Finding 3G Mobile Network Cells with Similar Radio Interface Quality Problems

Pekka Kumpulainen, Mika Särkioja, Mikko Kylväjä, Kimmo Hätönen

Finding 3g Mobile Network Cells with Similar Radio Interface Quality Problems

Pekka Kumpulainen[1], Mika Särkioja[2], Mikko Kylväjä[3], Kimmo Hätönen[4]

[1] Tampere University of Technology, Department of Automation Science and Engineering, Korkeakoulunkatu 3, 33720 Tampere, Finland
[2] Nokia Siemens Networks, BSO OSS Radio Network Optimizer, Espoo, Finland
[3] Aditro Software, Espoo, Finland
[4] Nokia Siemens Networks, CTO Research, Espoo, Finland
pekka.kumpulainen@tut.fi, {mika.1.sarkioja, kimmo.hatonen}@nsn.com, mikko.kylvaja@aditro.com

Abstract. A mobile network provides a continuous stream of data describing the performance of its cells. Most of the data describes cells with acceptable performance. Detecting and analysing mobile network cells with quality problems from the data stream is a tedious and continuous problem for network operators. Anomaly detection can be used to identify cells, whose performance deviates from the average and which are potentially having some sub-optimal configuration or are in some error condition. In this paper we provide two methods to detect such anomalously behaving cells. The first method estimates the distance from a cell to an optimal state and the second one is based on detecting the support of the data distribution using One-Class Support Vector Machine (OC-SVM). We use the methods to analyse a data sample from a live 3G network and compare the analysis results. We also show how clustering of found anomalies can be used to find similarly behaving cells that can benefit from the same corrective measures.

Keywords: Mobile network monitoring, radio interface quality, problem detection, OC-SVM

1 Introduction

Monitoring the mobile network is a crucial task for operators to maintain the service quality level and keep their subscribers satisfied. Efficient monitoring is needed to reduce the waste of network resources and to solve network performance related problems in a short time. The network produces massive amounts of data and a lot of experts' time is required for analysis. The majority of the available data are irrelevant in supporting troubleshooting and solving a particular network problem. Relevant observations need to be identified from the mass and presented for an expert. Therefore efficient tools are required to assist in the management of the network.

In 3G networks cells in the same layer are sharing the same frequency and air interface bandwidth is utilized more efficiently than in GSM networks. Therefore in

3G networks interference control is more important. 3G networks are designed to support a range of different services that have different behaviour and problems. [7]

The purpose of performance monitoring is to reveal problems in network elements, which have similar symptoms and therefore are more likely to require the same corrective actions. Analysis results produced by automated algorithms are independent of human operator's domain knowledge and experience. Therefore identification of problems can be automated and results obtained with greater average speed and accuracy. This helps in reducing the analysis time and implementing automated corrective actions.

In this paper we propose a methodology for detecting the most severe problems in the network and categorizing them into user friendly form that enable rapid and efficient actions in the network performance management. We compare two methods for detecting the problems. The first one is based on distances from ideal state of the network. It has been used in monitoring the radio interface quality measurements of GSM network [11]. This is enabled by the scaling method we use, which transforms the primarily unsupervised anomaly detection problem into a special case of semi supervised task [10, 11]. The second method we use is based on detecting the support of the distribution of the data using One-Class Support Vector Machine (OC-SVM) [14]. Finally, we cluster the detected anomalies into groups that present possible performance problems in the network.

First, in the following section, we introduce some measurements that are used for monitoring in daily routines. Next, we present a scaling procedure that utilizes the knowledge of the experts of the network operator. In section 4 we introduce two methods to detect possible problems in the network. We give examples of the behaviour of both methods using two quality variables for 2-dimensional visualization. We compare the results that these methods produce for quality variables of 3G network.

Finally, in Section 5, we cluster the detected problems to reveal problem groups and present descriptions given by network experts. Concluding remarks are given in the last section.

2 Radio Interface Performance Data

A huge amount of various events in the 3G network are counted and summed along time into counters. Although the counters contain all the information available about the network, it is too scattered to be of practical use. Therefore the counters are aggregated into Key Performance Indicators (KPI). They depict more understandable higher level information such as ratios of dropped calls or successful data connections. Each KPI is calculated from several counters using formulas, some of which are confidential.

In this study we use a limited set of radio network performance KPIs. We focus in clustering of call setup problems using the following KPI's:

- *Dropped call rate* (DCR) measures the proportion of initiated calls that are terminated abnormally.

- *HSDPA Setup Success Rate* indicates the proportion of successful HSDPA connection attempts. HSDPA is an enhancement for 3G network that enables fast mobile internet in downlink direction.
- *HSUPA Setup Success Rate* indicates the proportion of successful HSUPA connection attempts. HSUPA is similar to HSDPA, but in uplink direction.
- *R99 Setup Success Rate* measures the packet data connection attempt success ratio defined in basic 3G standard. HSUPA and HSDPA are newer technologies than R99 and they provide more radio interface capacity and throughput.
- *Voice call setup success rate* measures the ratio of successful voice call establishments.

We use a data set that is collected from a real functioning 3G network The KPIs are cell specific daily averages. The data consists of 123 cells from a period of 135 days. Due to missing values, the number of observations of individual cells varies from 120 to 135.

3 Scaling

Scaling the variables is an essential preprocessing procedure in practically all data mining tasks. Proper scaling is especially important in clustering as Gnanadesikan et al. [5], who refer to scaling as one method of weighting the variables, have pointed out: *When done efficiently, weighting and selection can dramatically facilitate cluster recovery. When not, unfortunately, even obvious cluster structure can be easily missed.*

We use a piecewise linear scaling procedure that utilizes the knowledge of the network experts [11]. The variables are scaled to interval [0, 1] so that 0 equals the worst and 1 equals the best possible performance as depicted in Fig. 1.

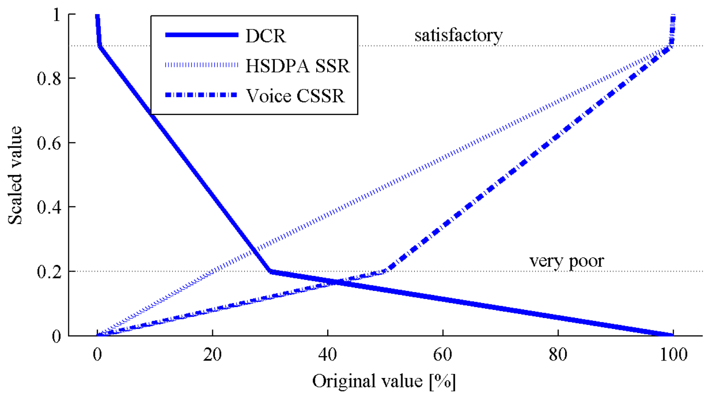

Fig. 1. Piecewise linear scaling of three radio performance KPIs.

A network management expert defines two corner points that correspond to very poor and satisfactory performance levels of each KPI. These corner points are scaled to appropriate values, 0.2 and 0.9 for example and the rest of the values are linearly interpolated. Scaling equalizes the importance of the variables in distance measures. This scaling method unifies the interpretation of the variables in the scaled space; the ideal performance will be 1 for all KPIs and the worst possible 0. Each value in the scaled space indicates the same level of performance for all KPIs.

4 Problem Detection

Mobile networks, as well as other industrial processes are controlled and mostly function in acceptable states of operation. Therefore the majority of the recorded performance data represent normal, acceptable behaviour. Possible problems show up as anomalous deviations from normal and anomaly detection methods can be used to reveal the problems.

We propose two methods for detecting problems in the performance of 3G mobile network. The first one is a simple distance based method that has been found working well in GSM network monitoring [10, 11]. The second one is based on OC-SVM which is an efficient tool for identifying the support of multivariate distributions [14].

4.1 Distance Based Detection

The scaling method maps the best performance values of all variables to 1. This converts the anomaly detection into a special semi supervised case. The ideal state is now known to consist of all ones. Thus, the distance from ideal can be used as a measure of total performance, larger distances representing more severe problems [10, 11]. Scaling each variable according to the severity of their performance level equalizes their effect in the distance calculation. An example of the distance based (DB) detection of two performance KPIs is presented in Fig. 2.

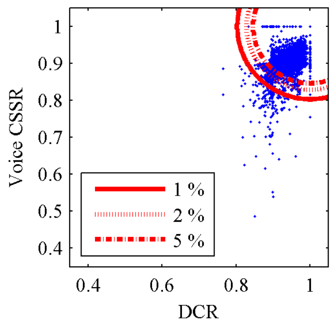

Fig. 2. Equal distance from ideal. 0.5%, 1% and 2% thresholds.

Equal Euclidean distance from ideal forms a circular area around the ideal. The threshold for detection can be set to a specific required distance or it can be identified from the data. This example displays percentage thresholds of the data.

4.2 One-Class support Vector Machine

Support Vector Machine (SVM) is a classifier that uses hypothesis space of linear functions in a high-dimensional kernel-induced feature space [3]. The basic SVM classifies a training vector $\mathbf{x}_i$ of l observations into two classes, defined by a class variable $y_i \in \{-1, 1\}$. Extensions, such as "one-against-one" have been developed for multiclass classification [8, 9].

One-class SVM (OC-SVM) was proposed by Schölkopf et al. for estimating the support of a high-dimensional distribution [14]. The strategy is to separate the data from the origin with maximum margin in the feature space induced by the kernel. Parameter $\nu \in (0, 1]$ controls the fraction of the observations that are allowed to be classified to the class of the origin. Those observations, lying outside the support of the distribution are considered as anomalies. A variety of nonlinear estimators in input space is achieved by using different kernel functions.

The maximum margin classifier to separate the data set from the origin is achieved by solving the following quadratic program:

$$\begin{aligned} \min_{\mathbf{w},\boldsymbol{\xi},\boldsymbol{\rho}} \quad & \frac{1}{2}\mathbf{w}^T\mathbf{w} - \rho + \frac{1}{\nu l}\sum_{i=1}^{l}\xi_i \\ \textit{subject to} \quad & \mathbf{w}^T\phi(\boldsymbol{x}_i) \geq \rho - \xi_i, \\ & \xi_i \geq 0, i = 1, \ldots, l . \end{aligned} \tag{1}$$

In this study we use a software package LIBSVM [1]. Its implementation of OC-SVM solves a scaled version of the dual form of (1):

$$\begin{aligned} \min_{\boldsymbol{\alpha}} \quad & \frac{1}{2}\boldsymbol{\alpha}^T\mathbf{Q}\boldsymbol{\alpha}, \\ \textit{subject to} \quad & 0 \leq \alpha_i \leq 1, i = 1, \ldots, l, \\ & \mathbf{e}^T\boldsymbol{\alpha} = \nu l, \mathbf{y}^T\boldsymbol{\alpha} = 0 . \end{aligned} \tag{2}$$

where $\mathbf{Q}$ is a positive semidefinite matrix $\mathbf{Q}_{ij} \equiv y_i y_j K(\mathbf{x}_i,\mathbf{x}_j)$ and the kernel $K(\mathbf{x}_i,\mathbf{x}_j) \equiv \phi(\mathbf{x}_i)^T\phi(\mathbf{x}_j)$. The kernel function can be selected freely. One commonly used and versatile [6] kernel is the Radial Basis Function (RBF) kernel: $K(\mathbf{x}_i,\mathbf{x}_j) = \exp(-\gamma\|\mathbf{x}_i - \mathbf{x}_j\|^2)$, $\gamma > 0$.

The decision function classifies the outliers to class -1 and the bulk of the data to class +1 by:

$$\mathrm{sgn}\left(\sum_{i=1}^{l}\alpha_i K(\boldsymbol{x}_i, \boldsymbol{x})\right) . \tag{3}$$

OC-SVM has been successfully used for document classification [12] and novelty detection for multi-channel combustion data [2]. Good results were achieved in both cases. However, both papers reported that the OC-SVM was very sensitive to the parameter γ of the RBF kernel function. As there is no classification information available, it is not possible to use methods based on error rates, such as cross validation, for parameter selection. The results have to be verified by the end users, application experts, and the parameters selected according to their subjective judgement. It should be noted that OC-SVM is not the only method that faces this challenge but it is reality with all unsupervised methods.

Fig. 3 presents anomaly thresholds detected from the same two KPIs as in earlier, now by OC-SVM with RBF kernel. In OC-SVM the ideal state (all 1) has no special meaning. In order to prevent the ideal cases from appearing as anomalies, the "definitely good enough" data points, within distance 0.1 from the ideal, have been removed. The parameter ν in (2) is adjusted so that the fraction of detected anomalies is 5% of the whole data set. When applied to these two KPIs from a 3G network the OC-SVM is not very sensitive to the kernel function parameter. Only minor changes are visible in Fig. 3 where two extreme values for the γ of the RBF kernel are used.

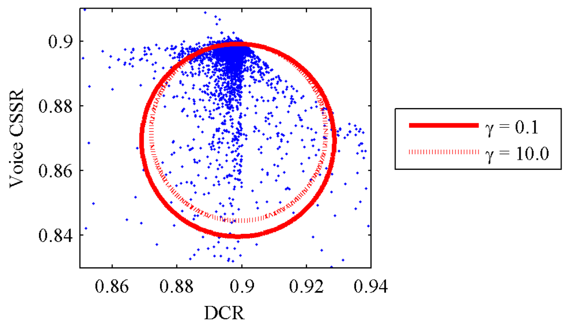

Fig. 3. Example of OC-SVM thresholds.

4.3 Results of Problem Detection

We applied both methods, DB and OC-SVM to a data set consisting of the 5 KPIs described in section 2. We set the threshold in the DB method to find the worst 5% of the data set, 816 observations. We applied OC-SVM to a subset of the data, where the data within 0.1 distance from **1** were removed as in the previous example. We used RBF kernel in OC-SVM with 5 values of γ and ν was 0.2061 which corresponds to 5% of the whole data set.

The number of detected anomalies for different γ varies between 831 and 832. There are 766 observations detected with all the γ values. Thus, for these date the OC-SVM does not seem to be too sensitive to the kernel function. However, only about half of the detected anomalies were common to those detected by the DB method. The numbers common observations are presented in Table 1.

Table 1. Number of common observations, detected by both methods.

Kernel parameter	$\gamma = 0.1$	$\gamma = 0.5$	$\gamma = 1$	$\gamma = 5$	$\gamma = 10$
Common with DB	412	415	419	442	463

The detected anomalies present potential problems in the network. The characteristics of the common and divergent observations are discussed in the following section where the problems are described by network experts.

5 Problem Categorization for Quality Monitoring

The end users of the monitoring applications appreciate simple but informative presentations of the results. Clustering is an effective way to summarize the information [4]. We use agglomerative hierarchical clustering with Ward linkage [15]. Hierarchical clustering requires the interpoint distances between the observations and therefore it is not suitable for large data sets. However, we are clustering a relatively small number of detected anomalies. One advantage is that the clustering is deterministic; there are no random elements as in the initialization of another popular clustering, k-.means [4].

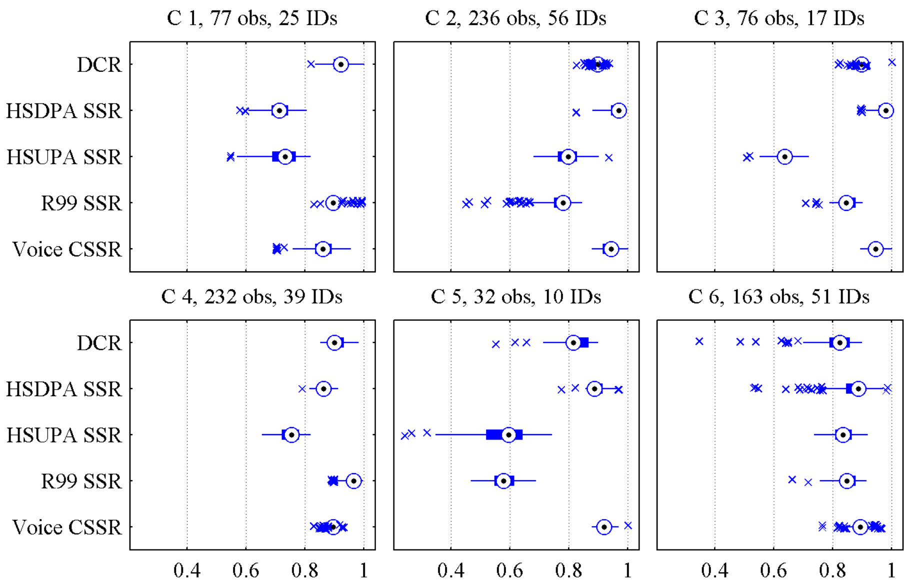

Fig. 4. Distance based problem clusters.

We use box plots [13] of the problem clusters produced to depict the characteristics of the clusters. The clustered anomalies detected by DB method are presented in Fig. 4 and those detected by OC-SVM in Fig. 5. Interpretation of the problem clusters is presented in Table 2.

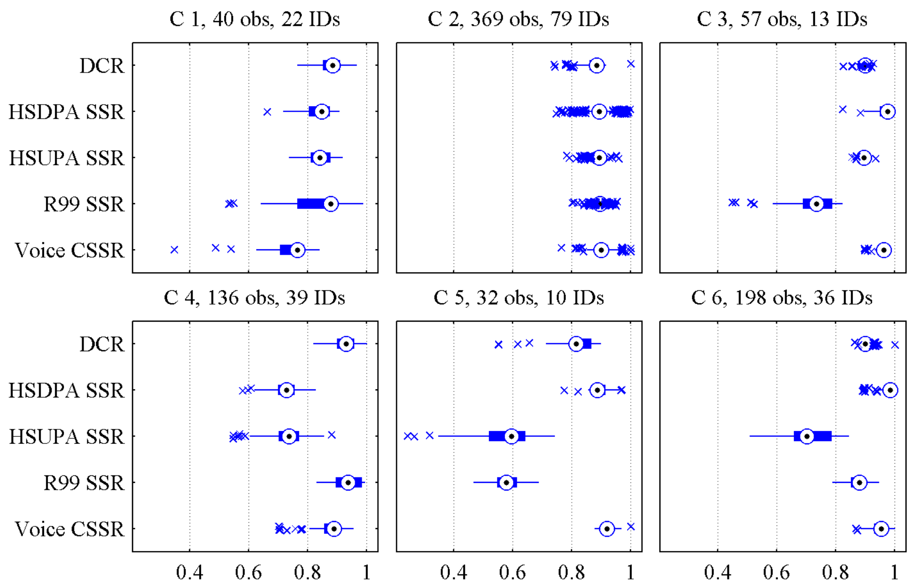

Fig. 5. OC-SVM problem clusters

Table 2. Descriptions of the problem clusters.

Problem	Distance based	OC-SVM
HSDPA channel conditions (bad quality because of cell overlap or no good dominant server)	Cluster 1 (77)	Cluster 4 (136)
Limited HSUPA license capacity, or limited channel elements available for HSUPA	Cluster 3 (76)	Cluster 6 (196)
Limited uplink coverage (probably because of indoor users.) Users make call attempts in bad radio conditions.	Cluster 5 (32)	Cluster 5 (32)
Radio access setup failures for R99 packet data connections, UL interference, configuration error or equipment failure.	Cluster 2 (236)	-
Channel element capacity limitation – Omniproblem. Probably heavy resource utilization	Cluster 6 (163)	Cluster 2 (369)
”Check HSUPA licenses and coverage map”	Cluster 4 (232)	-
“Analyze more detailed setup failure reason codes for Voice”. Behavior similar to Cluster 2		Cluster 1 (40)
Unexpected behavior, can be caused by particular mobile terminals or smart phones		Cluster 3 (57)

The identified clusters are mapped into known network problems based on symptoms. In general, DCR correlates with bad radio conditions, and in (Cluster 5/5) there are both setup and DCR problems. This is an indication of bad radio conditions. On the other hand, BTS unit capacity is one typical problem, as the system has limitations in handling a large number of connections simultaneously. In (Cluster 6/2) this can take place depending on the amount of traffic and capacity available. Independent of the physical hardware capacity there can be licensed capacity limitations; in (Cluster 3/6) this takes place in one direction, generating HSUPA connection setup problems. HSDPA does not benefit from soft handover like other channels and is more sensitive to channel conditions because of cell overlap, in (Cluster 1/4). R99 data traffic is used when a user requests small data packets or cannot have a HSPA connection, in case of Cluster (2/-) repeated R99 setup problems due to user location or system reservation failures. The voice connections share the same resource pool as HSUPA connections, but with higher priority. In normal conditions voice setup performance should be better than HSUPA. This indicates specific voice service problem in (Cluster -/1).

6 Conclusion

Monitoring the mobile network is a crucial task for operators in order to maintain the service quality level and keep their subscribers satisfied. Majority of the vast amounts of data is irrelevant for daily network performance management. In this paper we present two methods to detect cells that have potential problems in their performance. We use a scaling procedure that utilizes a priori expert knowledge and enables a unified interpretation of different performance KPIs.

We demonstrated how the information of the potential problems can be summarised for a radio expert to simplify daily performance monitoring routines. The detected problem clusters represent plausible groups of cells with similar problematic behaviour. Interpretation of the behaviour patterns requires an expert with sufficient knowledge of the network troubleshooting. However, with proper visualization an expert is able to give understandable explanations to the problems.

This method can be improved in future by introducing a problem library, where the most typical problem types are stored with suggestions of possible corrective actions. Additionally this method can be improved by collecting data and developing analysis method using data from several networks and different operators.

References

1. Chang, C-C., Lin, C-J.: LIBSVM: a library for support vector machines. http://www.csie.ntu.edu.tw/~cjlin/papers/libsvm.pdf (2001)
2. Clifton, L., Yin, H., Zhang, Y.: Support Vector Machine in Novelty Detection for Multi-channel Combustion Data. In: Wang, J., Yi, Z., Zurada, J., Lu, B-L., Yin, H. (eds) Advances in Neural Networks - ISNN 2006. LNCS 3973, pp. 836--843. Springer Berlin / Heidelberg (2006)

3. Cristianini, N., Shawe-Taylor, J.: An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press (2000)
4. Everitt, B., Landau, S., Leese, M.: Cluster analysis. Arnold, London (2001)
5. Gnanadesikan, R., Kettenring, J.R., Tsao, S.L.: Weighting and selection of variables for cluster analysis. Journal of Classification. 12 (1), 113--136 (1995)
6. Hsu, C-W., Chang, C-C., Lin, C-J.: A Practical Guide to Support Vector Classification, http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf
7. Kaaranen, H., Ahtiainen, A., Laitinen, L., Naghian, S., Niemi, V.: UMTS Networks. John Wiley & Sons, Chichester (2009)
8. Knerr, S., Personnaz, L., Dreyfus. G.: Single-layer learning revisited: a stepwise procedure for building and training a neural network. In: J. Fogelman, (ed), Neurocomputing: Algorithms, Architectures and Applications. Springer-Verlag, (1990)
9. Kreßel, U.: Pairwise classification and support vector machines. In: Schölkopf, B., Burges, C. J. C., Smola, A. J. (eds), Advances in Kernel Methods – Support Vector Learning, pp. 255--268, MIT Press. Cambridge, MA (1999)
10. Kumpulainen, P., Kylväjä, M., Hätönen, K: Importance of Scaling in Unsupervised Distance-Based Anomaly Detection. In: Proceedings of IMEKO XIX World Congress. Fundamental and Applied Metrology. September 6-11, 2009, Lisbon, Portugal. pp. 2411--2416 (2009)
11. Kylväjä, M., Kumpulainen, P., Hätönen, K.: Information Summarization for Network Performance Management, In: Laszlo M., Zsolt, J.V. (eds.) Proceedings of the 10th IMEKO TC10 International Conference on Technical Diagnostics, Budapest, Hungary, pp. 167-172 (2005)
12. Manevitz, L., Yousef, M.: One-class svms for document classification. The Journal of Machine Learning Research, 2, Special issue on kernel methods. 139--154 (2002)
13. McGill, R., Tukey, J.W., Larsen, W.A.: Variations of Boxplots. The American Statistician, 32, 12-16 (1978)
14. Schölkopf, B., Platt, J., Shawe-Taylor, J., Smola, A.J., Williamson, R.: Estimating the support of a high-dimensional distribution. Neural computation, 13 (7), pp. 1443-1472 (2001).
15. Ward, J.H. Jr.: Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association 58 (301) 236—244 (1963)