University of Groningen

Automatic fault detection and diagnosis in cellular networks using operations support systems data

Rezaei , Samira; Radmanesh, Hamidreza; Alavizadeh, Payam; Nikoofar, Hamidreza; Lahouti, Farshad

Published in:
2016 IEEE/IFIP Network Operations and Management Symposium (NOMS 2016)

DOI:
[10.1109/NOMS.2016.7502845](https://doi.org/10.1109/NOMS.2016.7502845)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2016

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Rezaei , S., Radmanesh, H., Alavizadeh, P., Nikoofar, H., & Lahouti, F. (2016). Automatic fault detection and diagnosis in cellular networks using operations support systems data. In *2016 IEEE/IFIP Network Operations and Management Symposium (NOMS 2016): Mini-Conference* (pp. 468-473). IEEE.
<https://doi.org/10.1109/NOMS.2016.7502845>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Automatic Fault Detection and Diagnosis in Cellular Networks Using Operations Support Systems Data

Samira Rezaei^{*}, Hamidreza Radmanesh⁺, Payam Alavizadeh⁺, Hamidreza Nikoofar⁺, Farshad Lahouti^{*}

^{*} School of Electrical and Computer Engineering, College of Engineering, University of Tehran

⁺ Mobile Communications Company of Iran

rezaei.samira@ut.ac.ir, [h.radmanesh, p.alavizadeh, hr.nikoofar]@mci.ir, lahouti@ut.ac.ir

Abstract—Self-Healing is one of the key important functionalities in self-organizing mobile communication networks. Despite its importance, self-healing has yet to receive a deserving research attention in the literature. This paper considers important quality indicators in a live mobile communications network and presents an automatic unified detection and diagnosis framework to identify root causes of faults occurred in the network. The proposed solution relies on unsupervised clustering of both traffic and signaling (continuous) key performance indicators for diagnosis. Hence, it is immune to the human error in the modeling phase. It also allows to effectively encapsulate experts knowledge as they relate clusters to root causes in the design phase. Detailed analysis of fault detection, the clustering schemes and the diagnosis are provided using operations support systems data of a real cellular network.

Keywords—self-healing; cellular network management; unsupervised learning; fault detection.

I. INTRODUCTION

Complexity of cellular networks is increasing rapidly. User behaviors affect the performance of these networks. Software or hardware problems, component breakdowns and harmful natural events also disturb quality of services in cellular networks. As a result an active self-healing system would be helpful in handling breakdowns. Although saving network from faults and failures is a crucial issue, there has been less attention to self-healing compared to self-configuration and self-optimization which are two other parts of self-organizing networks (SON). SON is identified as a key design principle for next generation networks [1]. It is also a fundamental element of LTE deployment in Release 8 onward [2]. As mentioned, three main functionalities of SONs are:

- a. Self-Configuration: Configuration of network elements such as base stations and eNodeBs are required in different phases including planning, extension and upgrade of network terminals. Due to the large number of nodes in the network, configuration of network elements is only feasible by self-configuration.
- b. Self-Optimization: Based on network situation and traffic, certain network parameters are to be recomputed and adjusted periodically. In legacy networks, optimization is done through periodic drive tests by the operators. Studies such as [3],[4] have already focused on providing self-

optimization solutions for networks.

- c. Self-Healing: Increasing complexity of cellular networks calls for an automatic process to handle cell outages or service degradations. A self-healing system has three main components: fault detection, root cause diagnosis and proposing corrective actions. Automatic root cause analysis is challenging in general and currently it is mainly carried out manually by the experts. This paper focuses on the two first stages of self-healing systems, i.e., automatic fault detection, and root cause diagnosis in cellular networks.

The number of studies on self-healing is limited because of two important reasons. First, the correspondence between common faults and their related symptoms is not well documented in the literature and is mainly an expert knowledge. Second, historical data of faults in mobile networks is often inaccessible to the scientific community. In addition, simulating these faults and modeling network behavior on simulators are often challenging and distanced from real complex cellular network settings [5]. Therefore, academia-industry collaboration is necessary to perform a thorough study of self-healing cellular systems. To the best of our knowledge, all previously reported studies have resorted to simulators and/or explored a limited set of faults and performance indicators [6]-[11].

An integrated detection and diagnosis framework is presented in [6]. Fault detection contains monitoring the radio measurements and comparing them to their normal behavior which are captured by their so-called profiles. Diagnosis of root causes is based on previous recorded fault cases and understanding their effects on the performance indicators. This study uses simulated data and considers three performance indicators (KPIs), namely channel quality indicator, call drop and handover timing advance.

The Bayesian classification in the context of supervised learning is considered to make a relation between faults and root causes in GSM and UMTS networks in [7]- [9]. In [7], it is confirmed that when sufficient historic data on faults and root causes are available relying on continuous key performance indicators for fault detection and diagnosis is advantageous. Indeed, discretized KPIs may only provide a

better performance when the training data set is insufficiently small. Both [9] and [8] use discretized KPIs. The focus in [9] is on ways of data driven model parameter learning to avoid the typical reliance of Bayesian approaches on experts' opinion. While both simulated and real data is used in [8], the identification of faulty cells relies only on one KPI (call drop rate).

To detect cell outage, a clustering algorithm, that utilizes four KPIs of reference and maximum-neighbor signal quality and power, is presented in [10]. The reported experiments are meant to detect two base stations with dropped transmission powers in a simulated LTE network 19 cells with 40 users in each cell.

In [11], a diagnosis method based on a supervised genetic fuzzy algorithm is presented. The genetic algorithm is used to learn the fuzzy rule base and as such relies on the existence of labeled training sets. The experiments are based on both a simulated data set and a real data set of 72 records with four KPIs and four root causes.

In this paper a unified framework for automatic fault detection and diagnosis in cellular networks is presented. The proposed framework exploits a high number of important continuous KPIs in operations support systems (OSS) data of a real mobile communication network. As a result we expect to recognize important and frequent faults in the network. The fault detection model is set up by discretizing network quality using KPIs of OSS data. For fault diagnosis, unsupervised clustering is used and it also incorporates the experts reasoning in the design to enable automatic decision support in an operating mobile communications environment. Details on faulty clusters, their validation, and the associated root causes are investigated and reported. In addition, the performance of different clustering algorithms within the framework is studied and compared. We use the real data of a live GSM network for performance evaluation. We anticipate that the proposed approach and framework can also be useful for other wireless standards in subsequent research and upon availability of data.

The rest of this article is organized as follows. Section II reviews the key performance indicators used in this work for cellular networks. Section III describes the proposed framework for automated fault detection and diagnosis and root cause analysis. In this section, the results obtained by using different unsupervised learning algorithms on dataset are also presented. Section IV concludes the article.

II. THE FRAMEWORK AND KPIs

Automatic troubleshooting of large scale cellular networks is vital to guarantee efficient utilization of network infrastructure and provisioning of users quality of experience. To this end, we consider the framework depicted in Figure 1 and analyze mobile network KPIs that are measured periodically across the network as OSS data. Due to the variety of measured KPIs and big amount of data, we resort to feature selection techniques utilizing experts' knowledge for an effective and more efficient analysis. As is the practice, records containing missing KPI values and outliers are removed from data. Faulty records are identified using network measurements of KPIs and categorized in clusters by pattern recognition techniques.

Next, taking the experts' knowledge into account, each cluster is associated with possible root causes for the fault and points towards actions for resolving them. In standard data mining design approaches, data cleaning usually precedes feature selection. In the current setting, since feature selection is based on experts' knowledge, one may opt to postpone data cleaning phase.

A. KPIs and Data Set

Many variables and counters measure every event in an operating mobile network. KPIs are computed based on these raw counters. Alarms also have a major role in fault detection of cellular networks and several studies have been conducted on alarm correlation to introduce different ways of interpreting multiple alarms. However, alarms are not sufficient for recognizing all potential faults and their causes in the network [12]. Besides alarms are mainly indicators of hardware problems and KPIs are more related to users' experiences.

Therefore, it is necessary to consider key performance indicators for monitoring the network performance. These KPIs are collected and averaged by network management systems periodically. Abnormal situations in the network affect the values of certain KPIs and this is how the experts in performance management and optimization team find out about the faults and take corresponding actions to resolve them. All considered KPIs in this paper are continuous variables. The distributions of the KPIs differ widely, and to better relate them in the proposed learning algorithms, min-max normalization is applied to the data set before analysis. The actual values corresponding to normalized values are used to interpret the results.

We divide the performance indicators of the mobile network to traffic and signaling categories. New incoming data is assessed in these two categories and possible faults are detected and diagnosed within the proposed framework. Details are reported in the subsequent Sections.

While the approach we take to tackle this problem is rather general, to arrive at concrete results and their validation, we use data from the operating GSM network of Mobile Communications Company of Iran (MCI, the biggest operator in the Middle East). Specifically, the dataset is from the operations support systems (OSS) data of a sub-network of MCI with 50 tri-sectorial sites in a dense urban environment over 30 days. The system provides average KPIs every hour. For each sector we consider 7 traffic hours so the number of records per day are 1050. The majority of KPIs are statistical and the number of efforts to access the network will affect them, so focusing on peak traffic hours enables more accurate design. In a similar direction in subsequent research, KPIs from other wireless networks can also be properly identified and taken into account in this framework. Below is a list of GSM KPIs we consider in the proposed fault diagnosis system.

1. Signaling channel related KPIs: A single mobile subscriber (MS) uses a Stand-alone Dedicated Control Channel (SDCCH) for establishing a call. Other usages of this channel are authentication, location updating and point to point SMS. Related KPIs to this channel are:

a. SDCCH Congestion Rate (SDCong): When reserving the SDCCH channel for a new inquiring MS fails due to congestion in the channel, the value of this KPI increases. Equipping more timeslots for SDCCH or modifying the Location Area Code (LAC) border in the case of high level of location updates are two possible solutions to reduce the load on SDCCH.

b. SDCCH Assign Success Rate (SDAssignSR): This KPI measures the percentage of successfully established SDCCH requests. Inappropriate values of this KPI may be due to interference in downlink signal, planning and hardware problems and high number of location updates.

c. SDCCH Drop Rate (SDDrop): The amount of dropped requests, which are successfully established on SDCCH channel is calculated by this KPI. The main reasons of high SDCCH drop may be low uplink and downlink signal quality.

2. Traffic channel related KPIs: Speech and data traffic are carried over GSM traffic channel (TCH). After establishing a SDCCH channel successfully, a TCH channel is assigned to the MS. Below is a list of TCH related KPIs.

a. TCH Congestion Rate (TCHCong): When there is not enough resource in the network to establish a call, congestion occurs. TCH congestion is one of the most important problems in the network and affects the quality immensely. One possible way to obviate this problem is adding new resources to the network, of course at a noticeable financial cost.

b. TCH Assignment Fail Rate (TCHAssignFR): It happens when the MS is not able to use the assigned TCH for voice calls. Lack of coverage, interference and hardware problems may affect this KPI to violate its normal behavior.

c. Signal Quality (RxQuality): It is the average quality of downlink and uplink signals. There are 8 levels in measuring signal quality; 0 is the best and 7 is the worst. Proportion samples of acceptable levels (0-4) shows the value of uplink and downlink signals quality. Many factors affecting the signal quality may have bad effects on other KPIs too.

d. Call Drop Rate (CDR): It indicates the percentage of not normally released channels to the number of all assigned channels to users for establishing voice calls.

III. FAULT DETECTION

Fault detection is the first step of any self-healing or decision support system for mobile network optimization. There are several ways to identify faults in a cellular network. Some studies used CDR as an important KPI for this purpose [7], [9]. This is while it certainly may not indicate all fault types.

Figure 1: Design of fault detection and diagnosis framework

Figure 2: Normalized histogram of CCC, two local optima $x=85$ and $x=94$ are indicators for discretizing network operation quality, Y axis is normalized frequency of data records with specified CCC.

In this paper, network quality is measured by Complete Correct Call (CCC) rate. The CCC is designed by *Sofrecom* (part of Orange group) and is used by *MCI* to measure quality of network operation as follows

$$CCC = 100 * (1 - SDCong/100) * SDAssignSR/100 * (1 - SDDrop/100) * (1 - TCHCong/100) * RxQuality * (1 - TCHAssignFR/100) * (1 - CDR) \quad (1)$$

The important KPIs according to experts' opinion are taken into account in this equation. To our knowledge, this includes the largest number of KPIs reported in the literature for fault detection so far. Consequently, the presented model can potentially detect a greater range of occurred faults in the network. The value of CCC is between 0 and 100. A higher value indicates a better network operation quality. The value of CCC can be discretized by the help of its distribution [13].

Figure 2 demonstrates the histogram of CCC in the considered dataset. There are two local optima in this diagram and the network operation quality can be discretized into three sections (excellent, acceptable and unacceptable). Our priority goes to cells which are in serious trouble and need immediate attention. According to Figure 2, records with CCC less than 85 are in this category and labeled as fault.

Data records are to be categorized into two classes of fault and not fault in order to benefit the classification algorithms for recognizing predictor KPIs. In fact the classification schemes can be used as an alternate approach to detect faults with only a limited subset of KPIs and with very good accuracy. The dataset is split into train (70%) and test (30%) data sections. We examined five classification algorithms, namely Chi-squared automatic interaction detection (CHAID), quick unbiased efficient statistical tree (QUEST), Bayesian network, support vector machine (SVM) and classification and regression tree (CRT) [14], [15] for this purpose.

Table 1: The accuracy of fault detection for different classification algorithms.

Method	CHAID	QUEST	Bayesian Network	SVM	CRT
Accuracy	92.1	93.66	90.57	94.2	93.11

Table 2: Elicited rules by QUEST. F1: SDCCH Assignment Success Rate, F2: SDCCH Drop and F3: Call Drop Rate.

SDAssignSR	SDDrop	CDR	Fault
≤ 86.1	-	-	yes
> 86.1 and ≤ 92.9	≤ 1.0005	-	no
> 86.1 and ≤ 92.9	> 1.0005	-	yes
> 92.9	-	≤ 0.5336	no
> 92.9 and ≤ 94.2	-	> 0.5336	yes
> 94.2	-	> 0.5336	no

Table 1 presents the accuracy of fault detection using the classification algorithms on the dataset. The SVM and QUEST perform closely well in the fault detection accuracy.

The QUEST is a binary-split decision tree and has the ability to explain results in an understandable way. It is also capable of recognizing the most considerable KPIs in producing faults. QUEST splits the whole data just by three KPIs and with the accuracy of 93.66%. These three KPIs affect network quality the most. Elicited rules are illustrated in Table 2. These rules in fact cover the ranges of all KPIs. As KPIs are continuous, each one can take place in several depths of the tree. Total depth of the tree is three and is quite small; hence, there is no over-fitting problem in decision trees.

IV. ROOT CAUSE ANALYSIS

By fault detection, normal records of data are excluded. The dataset for following experiments contains faulty records and by faulty records we mean the records with the value of CCC lower than 85 (according to Figure 2). The purpose of this paper is to distinguish different fault types in a cellular network and identify related root causes for each one. To this end we apply unsupervised learning algorithms including expectation maximization (EM), density-based spatial clustering of applications with noise (DBSCAN), agglomerative hierarchical clustering, Xmean and Kmeans [16] to the faulty records. The experiments show that among the clustering algorithms studied, EM has the best performance and can find meaningful clusters. Each cluster indicates a specific fault and experts are requested to assign a root cause to each of them.

The purpose in EM is to maximize the exception of $\ln p(D_{unobserved}|\theta)$. EM starts with θ as an initial estimation of maximum likelihood and improves it iteratively to increase the likelihood of observed data. It has two main steps after initializing the θ parameter. If $|\theta^{t+1} - \theta^t| < \epsilon$, the algorithm stops, otherwise it continues to the expectation step. In the expectation step, EM uses the current parameters estimation of observed data and computes

$$Q(\theta; \theta^t) = E_{unobserved}(\ln p(D_{unobserved}|\theta)|D_y, \theta^t) \quad (2)$$

$\ln p(D_{unobserved}|\theta)$ is a linear function of the unobserved variable and D_y is the observed data. In the next step, the expectation of last prediction is maximized by:

$$\theta^{t+1} = \arg \max_{\theta} Q(\theta; \theta^t) \quad (3)$$

In dealing with a new incoming OSS data record, the value of CCC first determines either a faulty or a normal operation quality. If the new record is labeled as a fault, then the closest cluster center to the KPIs of this record indicates the associated root cause. It is important to use a cluster validation method. There is no gold data for a big real data in this context. As a result Silhouette coefficient [17] is used as an internal method to validate the obtained clusters. It is based on cohesion and separation of clusters and is also capable of finding the optimum number of clusters. The Silhouette coefficient is computed for each point o in cluster C_i as follows

$$S_o = \frac{(b_o - a_o)}{\max(a_o, b_o)} \quad (4)$$

$$b_o = \min_{\substack{C_j: 1 \leq j \leq N \\ i \neq j}} \left\{ \frac{\sum_{o' \in C_j} \text{dis}(o, o')}{|C_j|} \right\} \quad (5)$$

$$a_o = \frac{\sum_{o' \in C_i, o' \neq o} \text{dis}(o, o')}{|C_i| - 1} \quad (6)$$

where b_o quantifies clusters separation and a_o is an indicator of cluster cohesion; N is the total number of clusters; $|C_j|$ is the number of records in cluster C_j and $\text{dis}(o, o')$ is the Euclidean distance between data points o and o' . The Silhouette coefficient is between -1 and 1. The closer the value to 1, the better the clustering of data would be.

A. Traffic Channel KPIs

Figure 3 shows the Silhouette coefficient averaged over the TCH data records for different clustering algorithms. It confirms that EM with 5 clusters has the best performance among all other algorithms.

Figure 4 depicts the distribution of TCH related KPIs (traffic category) in four faulty clusters. Note that one of the five clusters in Figure 3 corresponds to normal TCH situation and is not illustrated in Figure 4. In the sequel, we analyze the TCH KPIs in different clusters and assess the root causes with expert knowledge.

Figure 4.a presents faulty records with a normal CDR. Instead it suffers from high value of congestion and failures in TCH assignments. Low values of signal quality confirm difficulties in TCH assignments. Congestion may be alleviated by enabling half rate feature in traffic channel (half rate refers to a system in which the codec operates at half the rate at a toll on quality).

It is probably better to use dynamic half rate switching threshold and tune the related parameters based on traffic load. In fact, the parameters are thresholds on when the network allocates full rate channel to users and when it allocates the half rate one. High path loss, high TCH blocking and high value of connection losses are the root causes in this cluster.

Figure 3: Averaged Silhouette coefficient of different clustering algorithms on TCH data

Figure 5: Averaged Silhouette coefficient of different clustering algorithms averaged over SDCCH data records.

Figure 4: KPIs distribution in four different faulty clusters of TCH data. Y axis is frequency of faulty records in each category. The curve with diamond, star, circle and plus markers indicate CDR, TCHCong, TCHAssignFR and RxQuality respectively.

Figure 4.b presents a cluster with congestion problem. It also suffers from high value of CDR and TCHAssignFR and signal quality. According to experts' opinion, the main root cause of faults in this cluster is hardware problem in base station controller (BSC) links or possibly problems in BSC transcoders.

Figure 4.c is a cluster with almost normal values for congestion indicator and TCHAssignFR. CDR is high and signal quality is inappropriate. Low value of signal quality may cause high value of bit error rate (BER). It also may cause more failures in handover requests. Interference problems and/or low signal strength in areas with no dominant server are the main reasons of high CDR due to bad quality. Missing neighbor relations could be another reason for bad quality because the mobile user is not connected to the strongest server and therefore perceives bad quality. One possible way is to investigate interfered cell in order to find the source of interference. Applicable GSM features such as frequency hopping and base station power control are other ways of reducing interference. Reasons for high values of CDR in Figure 4.d are the same as those in Figure 4.c, except that there are now problems in TCH assignments. Failures in TCH assignments especially in urban areas are because of hardware problems on TRXs, feeders and antennas.

B. Signaling Channel KPIs

Experiments for optimum number of clusters for signaling data is shown in Figure 5. The Silhouette coefficient for EM with 6 clusters is 0.503. Figure 6 shows the distribution of KPIs in these clusters.

Figure 6.a represents normal situation of signaling KPIs, therefore, these records have some problems in their traffic related KPIs levels. Figure 6.b indicates high value of SDCCH congestion. Two other KPIs are at their normal levels in this cluster. Defining more SDCCH channel or enabling dynamic SDCCH feature are two possible solutions to overcome the congestion issue.

Figure 6.c has both SDCong and SDDrop issues. These two KPIs are too important as their low values affect the user experience adversely. In this cluster, the value of CDR should be checked. The reasons of abnormal values for SDDrop in this case is similar with the high CDR. Some environmental factors like coverage, interference and poor transmission quality affect the value of SDDrop.

Congestion is a critical issue in Figure 6.d. SDAssignSR is also unacceptable and SDDrop is not admissible. Consequently these cells suffer from weak coverage area, and radio signal strength. It is also possible to have interferences, such as intra - network interference, improper frequency planning or other external interference. To reduce congestion it is recommended to check LAC border or decrease the signaling load on the related cells. Figure 6.e is similar to Figure 6.d in both SDDrop and SDAssignSR but it has no problem in congestion. Figure 6.f has hardware problem because of very bad values of SDAssignSR. Unstable transmission links over the Abis interface (Abis interface is responsible for transmitting traffic and signaling information between the GSM BSC and the GSM base transceiver station) is the possible root cause for this cluster.

Figure 6: KPIs distribution in six different clusters of SDCCH data records. Y axis is frequency of faulty records in each category. The curve with diamond, star and circle markers indicate SDDrop, SDCong and SDAssignSR, respectively.

V. CONCLUSIONS

Automatic troubleshooting of large scale cellular networks is vital to guarantee efficient utilization of network infrastructure and provisioning of users quality of experience. The complexity of managing these networks and big amount of produced data, persuade operators to benefit self-healing systems. In this paper a unified fault detection and diagnosis framework has been presented. The presented framework utilizes unsupervised learning on continuous signal and traffic KPIs and is designed to capture the expert's knowledge on root cause analysis effectively. Our results indicate that three specific KPIs, when analyzed effectively, can flag almost all faulty OSS records in a GSM network. The proposed framework serves as a decision support system for experts towards efficient resolution of cellular network faults. We note that while there are differences among different technologies and equipment manufacturers on their KPIs and how they are measured, the proposed approach and framework still applies and can accommodate possible differences.

REFERENCES

- [1] N. Alliance, "Recommendation on SON & O&M." 2008.
- [2] 3GPP TS 32.541, "Self-Organizing Networks (SON); Self-Healing Concepts and Requirements".
- [3] T. F. and S. Nishioka, "Selective handover for traffic balance in mobile radio communications," *IEEE International Conference on Communications*, vol. vol.4, pp. 1840–1846.
- [4] S. Kim and P. Varshney, "Adaptive load balancing with preemption for multimedia cellular networks," *IEEE Wireless Communications and Networking*, vol. 3, pp. 1680–1684, 2003.
- [5] R. Barco, P. Lazaro, and P. Munoz, "A unified framework for self-healing in wireless networks," *Communications Magazine, IEEE*, vol. 50, no. 12, pp. 134–142, 2012.

- [6] I. Transactions and O. N. Network, "An Automatic Detection and Diagnosis Framework for Mobile Communication Systems," *IEEE Transactions on Network and Service Management*, vol. 9, no. 2, pp. 184–197, 2012.
- [7] R. B. et Al, "Continuous Versus Discrete Model in Autodiagnosis Systems for Wireless Networks," *IEEE Trans. Mobile Computing*, vol. 7, pp. 673–81, 2008.
- [8] R. M. Khanafer, B. Solana, J. Triola, R. Barco, L. Moltsen, Z. Altman, and P. Lazaro, "Automated Diagnosis for UMTS Networks Using Bayesian Network Approach," *Vehicular Technology, IEEE Transactions on*, vol. 57, no. 4, pp. 2451–2461, 2008.
- [9] R. Barco, V. Wille, L. Díez, and M. Toril, "Learning of model parameters for fault diagnosis in wireless networks," *Wirel. Netw.*, vol. 16, no. 1, pp. 255–271, Jan. 2010.
- [10] Yu Ma, M. Peng, W. Xue, and X. Ji, "A Dynamic Affinity Propagation Clustering Algorithm for Cell Outage Detection in Self-Healing Networks," in *IEEE Wireless Communications and Networking*, 2013.
- [11] E. J. Khatib, R. Barco, A. Gomez Andrades, and I. Serrano, "Diagnosis based on genetic fuzzy algorithms for LTE Self-Healing," *Vehicular Technology, IEEE Transactions on*, vol. PP, no. 99, p. 1, 2015.
- [12] R. Barco, V. Wille, and L. Díez, "System for automated diagnosis in cellular networks based on performance indicators," *European Transactions on Telecommunications*, vol. 16, no. 5, pp. 399–409, 2005.
- [13] Pekko Vehviläinen, "Data Mining for Managing Intrinsic Quality of Service in Digital Mobile Telecommunications Networks," Tampere University of Technology, 2004.
- [14] W.-Y. Loh, "Classification and regression trees," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 1, no. 1, 2011.
- [15] T. N. Phyu, "Survey of Classification Techniques in Data Mining," in *Proceedings of the International MultiConference of Engineers and Computer Scientists*.
- [16] Lior Rokach, "A survey of Clustering Algorithms," in *Data Mining and Knowledge Discovery Handbook*, 2010, pp. 269–298.
- [17] P. J. Rousseeuw, "Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis," in *Computational and Applied Mathematics*, 1987, pp. 53–65.