Hindawi Publishing Corporation
Journal of Computer Networks and Communications
Volume 2012, Article ID 163184, 13 pages
doi:10.1155/2012/163184

Research Article

An Approach for Network Outage Detection from Drive-Testing Databases

Jussi Turkka,[1,2] Fedor Chernogorov,[2] Kimmo Brigatti,[2] Tapani Ristaniemi,[2] and Jukka Lempiäinen[1]

[1] *Department of Communications Engineering, Tampere University of Technology, 33720 Tampere, Finland*
[2] *Department of Mathematical Information Technology, University of Jyväskylä, 40014 Jyväskylä, Finland*

Correspondence should be addressed to Jussi Turkka, jussi.turkka@tut.fi

Received 18 March 2012; Revised 24 September 2012; Accepted 25 September 2012

Academic Editor: Sayandev Mukherjee

A data-mining framework for analyzing a cellular network drive testing database is described in this paper. The presented method is designed to detect sleeping base stations, network outage, and change of the dominance areas in a cognitive and self-organizing manner. The essence of the method is to find similarities between periodical network measurements and previously known outage data. For this purpose, diffusion maps dimensionality reduction and nearest neighbor data classification methods are utilized. The method is cognitive because it requires training data for the outage detection. In addition, the method is autonomous because it uses minimization of drive testing (MDT) functionality to gather the training and testing data. Motivation of classifying MDT measurement reports to periodical, handover, and outage categories is to detect areas where periodical reports start to become similar to the outage samples. Moreover, these areas are associated with estimated dominance areas to detected sleeping base stations. In the studied verification case, measurement classification results in an increase of the amount of samples which can be used for detection of performance degradations, and consequently, makes the outage detection faster and more reliable.

1. Introduction

Modern radio access networks (RAN) are complex infrastructures consisting of several overlaying and cooperating networks such as next-generation high-speed-packet-access (HSPA) and long-term evolution (LTE) networks and as such are prone to the impacts of uncertainty on system management and stability. Classical network management is based on a design principle which requires knowledge of the state of all existing entities within the network at all times. This approach has been successfully applied to networks of limited scale but it is foreseen to be insufficient in the management of future complex networks. In order to maintain a massive multivendor and multi-RAN infrastructure in a cost-efficient manner, operators have to employ automated solutions to optimize the most difficult and time-consuming network operation procedures. Self-organizing network concept [1] has emerged in the last years, with the goal to foster automation and to reduce human involvement in management tasks. It implies autonomous configuration, optimization, and healing actions which would result in a reduced operational burden and improve the experienced end user quality-of-service (QoS). One of the downsides of the SON concept is the necessity to gather larger amounts of operational data from user equipment (UE) and different network elements (NE).

To guarantee sufficient coverage and QoS for subscribers in indoor and outdoor environments, mobile operators need to carry out various radio coverage measurements. In the past, manual drive tests have been employed for this purpose. However, there are some challenges and limitations in manual drive testing that could be improved. Firstly, manual drive testing is a resource-consuming task requiring a lot of time, specialized equipment, and the involvement of highly qualified engineers. Secondly, it is impossible to capture the full coverage data from every geographical location by using manual drive testing, since most of the UE generated traffic comes from indoor locations, while drive testing is limited mainly to roads. The cost and reachability limitations of manual drive testing prompts the research towards automated UE-assisted data gathering solutions which can minimize the need for manual drive testing and allow gathering of more comprehensive databases. If UEs measure the radio coverage periodically and provide the measurements together with location and time information to the network, then large radio environment databases with user-perceived coverage experience can be built to support the RAN operation and optimization. However, essential problems with these large databases are the information overflow and a "curse" of dimensionality. Those problems need to be addressed while analyzing and transforming the raw measurement data in these huge operational databases into meaningful information. This paper describes an approach to the above-mentioned problems by proposing a data-mining framework for the analysis of the UE-reported radio measurements. This approach allows the detection of the coverage problems in a cellular network on the basis of learning the network's prior operational behavior. The proposed framework is validated with simulations by using *Renesas Mobile Europe's* state-of-art LTE system simulator to construct large MDT measurement databases.

The article is organized as follows: Section 2 describes the Minimization of Drive Tests concept which can be used to gather and build the UE measurement report databases for HSPA and LTE networks with the focus on coverage aspects. Section 3 describes the data-mining framework which is used for the analysis of the MDT databases, and finally, Section 4 describes simulation scenarios and the performance evaluation results of outage detection caused by a specific type of network failure known as "*sleeping cell.*"

2. Minimization of Drive Tests

Minimization of Drive Testing use cases for self-organizing networks were introduced by the operators alliance Next Generation Mobile Networks (NGMN) during 2008 [2] and at the time of writing this, the MDT solutions are researched by the network vendors and operators in the 3rd Generation Partnership Project (3GPP) [3, 4]. The goal of the MDT research in 3GPP is to define a set of measurements, measurement reporting principles and procedures which would help to collect coverage-related information from UEs. MDT feasibility study phase [3] started at late 2009 and during 2010 it focused on defining the reported measurement entities and MDT use cases for example, coverage optimization and QoS verification. Coverage optimization use case targets for the detection of such network problems as coverage holes, weak coverage, pilot pollution, overshoot coverage, and issues with uplink coverage, as described in [3]. After the feasibility study, the research focused on defining MDT measurement, reporting and configuration schemes for LTE release 10 during 2011 [4]. The MDT measurement and reporting schemes are *immediate MDT* and *logged MDT*. The immediate MDT scheme extends Radio Resource Control (RRC) measurement reporting to include the available location information to the measurement reports for UEs which are in connected mode [4]. In the logged MDT scheme, the UEs can be configured to collect measurements in idle mode and report the logged data to the network later [4]. After the release 10, the main focus of MDT work will be on enhancements in the availability of the detailed location information and improvements in QoS verification [5].

2.1. MDT Measurement Configuration. MDT measurements can be configured in LTE either by using management based or signaling-based configuration procedures [4, 6]. In the *management-based configuration*, the base stations are responsible for configuring all selected UEs in a particular area to do the immediate or logged MDT measurements [4, 6]. The *signaling-based MDT* is an enhancement to a signaling-based *subscriber and equipment trace* functionality [6] where the MDT data is collected from one specific UE instead of a set of UEs in a particular area. Detailed signaling flows for activating MDT measurements are described in [6].

The MDT measurement functionality allows operators to collect measurements either periodically or at an instance of a trigger such as a network event [3, 4]. The measurement report consists of the available location, time, cell-identification data and radio-measurement data. There are different mechanisms for estimation of user locations. The most coarse location info is the serving Cell Global Identification (CGI) and in the best case the detailed location is obtained from the Global Navigation Satellite System (GNSS). The cell identification info consists of the serving-cell CGI or Physical Cell Identifications (PCI) of the detected neighboring cells. The radio measurements for the serving and neighboring cells include the reference signal received power (RSRP) and reference signal received quality (RSRQ) for LTE system and common pilot channel received signal code power (RSCP) and received signal quality (E_C/N_0) for HSPA system [3, 4].

2.2. Logged MDT. The logged MDT measurement and reporting scheme enables data gathering from the UEs which are camped normally in RRC idle state. The logged MDT configuration is provided to the UEs via RRC signaling while UEs are in RRC connected mode. Logged mode configuration parameters are listed and described with more details in [4]. When the UE moves to the RRC idle state, MDT measurement data that is, time, location info and radio measurements, are logged to UE memory. The network can ask UEs to report the logged data when UE returns back to RRC connected state. Currently there can be only one RAT specific logged MDT configuration per UE which is valid only for the RAN providing the configuration. If an earlier configuration exists it will be replaced by newer one [4]. Since the logged MDT mode is an optional feature for UEs, this paper focuses more on the immediate MDT which will be a tool for operators to gather the measurements from LTE release 10 and onwards.

2.3. Immediate MDT. Immediate MDT is based on the existing RRC measurement procedure with an extension to include the available location information to the measurement reports. LTE release 10 RRC specifications [7] allow

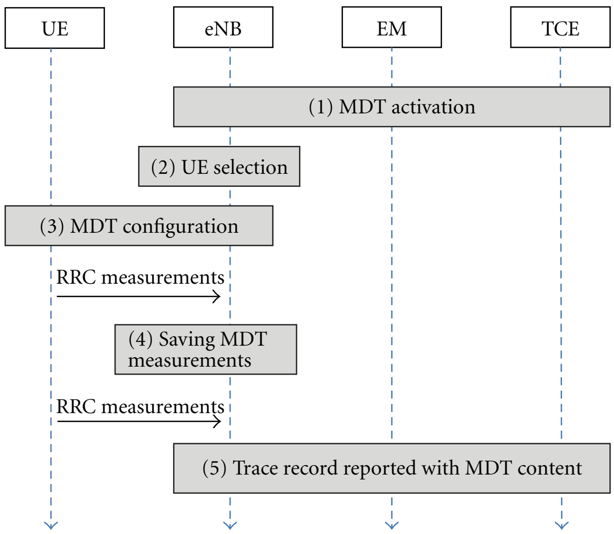


Figure 1: Immediate MDT reporting.

operators to configure RRC measurements in a way that RSRP and RSRQ measurements are reported periodically from the serving cell and intrafrequency, interfrequency and inter-RAT neighboring cells with the available location information. The immediate MDT measurement reporting principles are depicted in Figure 1 as described in [6].

Before the immediate MDT reporting can be started, a base station—E-UTRAN NodeB (eNB) is activated and configured to collect immediate MDT measurements. In step 1, an element manager (EM) sends a cell trace session activation request to the eNB including MDT configuration so that the eNB can later report the trace records back to the trace element (TCE). After the cell traffic trace activation, the eNB selects the UEs for MDT while taking into account the user consent that is, users permission for an operator to collect the MDT measurements. The eNB sends the RRC measurement configurations to the selected UEs for example, reporting triggers, intervals, and list of intrafrequency, interfrequency and inter-RAT measurements with a requirement that UEs include the available location information into the measurement reports as specified in the RRC specification information element (IE) *ReportConfigEUTRA* field [7].

When the RRC measurement condition is fulfilled for example, a periodical timer expires or a certain network event occurs, the UE sends available RSRP and RSRQ measurements to the eNB with the available *LocationInfo* IE added to the measurement report [7]. If detailed location information is available, then the latitude and the longitude are included into the measurement report. If the detailed location information is obtained by using GNSS positioning method then the UE shall attach time information to the report as well [4]. This GNSS time information is used to validate the detailed location information. Note that in case of the immediate MDT, the UE does not send the absolute time information as it does in case of logged MDT. The eNB is responsible for adding the time stamp to the received MDT measurement reports when saving the measurements to the trace record.

2.4. MDT Database. The MDT database is constructed by collecting the MDT measurements from the network. In our study the MDT database consists of periodical measurements, as well as measurements collected at the time instance of A3 (A3 event is E-UTRAN RRC measurement event which triggers when neighboring cell becomes an offset better than the serving cell) events preceding successful intra-LTE handovers (HO) and radio link failures (RLF). It is assumed that each measurement sample in the analyzed database consists of 22 features as described in Table 1.

The MDT measurement samples consist of the latitude, longitude, serving cell, and neighboring cell radio measurements reported by the UE. In addition, time information, serving-cell wideband channel quality indicator (WCQI) and uplink power headroom report (PHR) values were added by the eNB. Moreover, a label of the report condition is always appended to a measurement sample, that is, eNB knows if the MDT data sample is a periodical, A3 event-triggered measurement report or UE RLF report [4]. Currently, the release 10 MDT specifications do not support the feature of collecting detailed location for A3 events. However, this feature is to be included to MDT in release 11. Therefore, the structure of the MDT measurement sample described in Table 1 is assumed to be common for all of these three types of MDT reports.

3. Outage Detection Data-Mining Framework

It is known that the SON framework includes three functionalities, namely self-configuration, self-optimization, and self-healing. Self-configuration is related to the initial steps of the network setup. Self-optimization is concentrated on monitoring the network state and automatic parameter tuning for achievement of the highest possible network performance without compromising the robustness of its operation. In case of a network failure or malfunction, the self-healing tries to autonomously detect problems, diagnose root causes, and compensate or recover from the malfunctioning state back to normal operation. A good example of self-healing is the cell-outage management [8, 9] use case in LTE networks, which aims to improve the offline coverage optimization process by detecting and mitigating outage situations automatically. For this purpose, the self-healing algorithm requires several key performance indicator (KPI) measurements from both eNBs and UEs. The KPIs such as cell load, RLF counters, handover failure rate or, UEs neighboring cell RSRP measurements may be used as indicators of the network outage [8]. In [9], the condition for the outage was based on predefined thresholds of received signal strength and quality. However, deployment of several self-organizing functionalities can increase significantly the number of measured and reported KPIs thus increasing the complexity of the network and SON architectures. This may result in new challenges for network engineers as well. Firstly, high-dimensional KPI databases of network measurements

Table 1: Structure of the MDT measurement.

Feature No.	Feature	Description
1	Time	Time stamp
2-3	Location	Latitude and longitude
4	Serving-cell info	CGI
5	RSRP	Serving-cell RSRP in dBm
6	RSRQ	Serving-cell RSRQ in dB
7–13	Three strongest intra-LTE neighbors	CGI and RSRP for each neighbor
14–20	Three best quality intra-LTE neighbors	CGI and RSRQ for each neighbor
21	Serving-cell wideband CQI	Indicator of wideband signal quality
22	Power headroom report	Available uplink transmission power

are created, making expert-driven manual data analysis for identifying the right KPI/fault-associations a complicated task. The KPI/fault-associations are needed for developing good algorithms. Secondly, since the networks are complex and dynamic in nature, it is not obvious which KPIs should be measured and how often. For example, how to select from among several performance indicators, those which are going to reveal a certain feature of the network behavior in the most meaningful and effective manner?

It is envisioned that the above-mentioned challenges can be solved with advanced machine learning and data-mining algorithms which rely on autonomous learning of network behavior and efficient processing of the high dimensional databases consisting of wide range of KPIs. The data-mining can be used for extracting interesting, previously unknown and potentially useful information patterns from the large databases [10]. Usually the data mining process consists of several phases such as data cleaning, database integration, task relevant data selection, data mining, and data-pattern evaluation [10]. Data cleaning, integration, and selection are data preprocessing phases where data is prepared for further analysis [10, 11]. The data mining itself can consist of several different functionalities such as classification of data, association of data, clustering of data, dimensionality reduction, and anomaly detection [10]. In the pattern evaluation phase, the information patterns are visualized and analyzed to see if novel and valid information can be extracted from them. Even if interesting information patterns are discovered, it does not mean that it is automatically usable or useful from the data mining problem point of view, and therefore, information patterns need to be validated.

Within the family of cell-outage use cases included into self-healing of cellular radio networks there is a specific problem called *sleeping cell*. The sleeping cell is a compound term, which includes erroneous network behavior ranging from performance degradation to complete service unavailability. A specific characteristic of sleeping cell is that the network performance is degraded but this degradation is not easily visible to network operators and thus detection of this problem with traditional alarming systems is a complicated and slow process as described in [12]. There is no definition of a certain network failure which would cause appearance of a sleeping cell, as there can be several reasons. One type of sleeping cell could be malfunction of eNB RF unit where the eNB transmission and reception capabilities degrade slowly to a point where transmission, reception, or both are not working anymore. This results in an outage situation where eNB cannot provide service for the UEs in the coverage area of the sleeping cell. Indicators which could reveal sleeping cells are degradation in handover activity, low call setup rates and low cell loading. Different kinds of indicators are needed to detect sleeping cells in live networks since networks consist of several overlapping frequency layers and radio access technologies. In [13], a sleeping cell is detected by using statistical classification techniques for graphs constructed from UE reported neighboring cell patterns. Changes in the neighboring cell patterns are used as indicators of outage.

One of the main goals of the research into the minimization of drive test is the development of algorithms which make operation of the networks more robust and efficient, so we developed a data-mining framework which detects coverage problems, such as sleeping cells, by using the high-dimensional MDT measurement databases. The data-mining framework described in this paper relies on dimensionality reduction which allows simplifying the anomaly detection and data classification processes. Motivation of using the dimensionality reduction is to make the framework robust and easily extendable with new numerical KPIs. On the other hand, the motivation of classifying MDT measurement reports to periodical, handover, and outage categories is to detect areas where periodical reports collected from certain frequency layer starts to show assumptions of outage. It is worth of noting that periodical MDT measurements can be collected from intra- and interfrequency layers simultaneously [4]. Therefore, some measurements for classification are available even if UE is connected on different frequency layer than the sleeping cell. This can happen in live networks where operators have deployed several overlapping frequency layers for capacity and coverage. If UE starts to experience outage on one frequency layer then it is handed over to another frequency layer before radio link failure occurs.

3.1. Data Mining Framework. The data-mining framework consists of learning and problem-detection phases. In the learning phase, the MDT database is constructed by collecting UE reported measurements from the network as depicted in Figure 2.

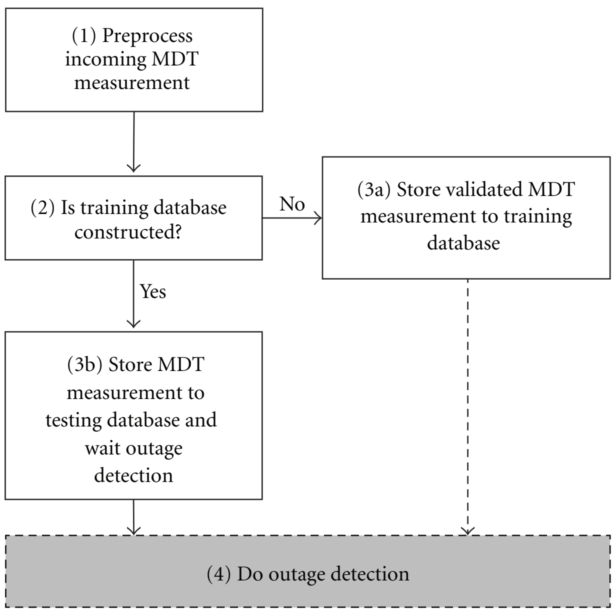


Figure 2: Data-mining learning phase description.

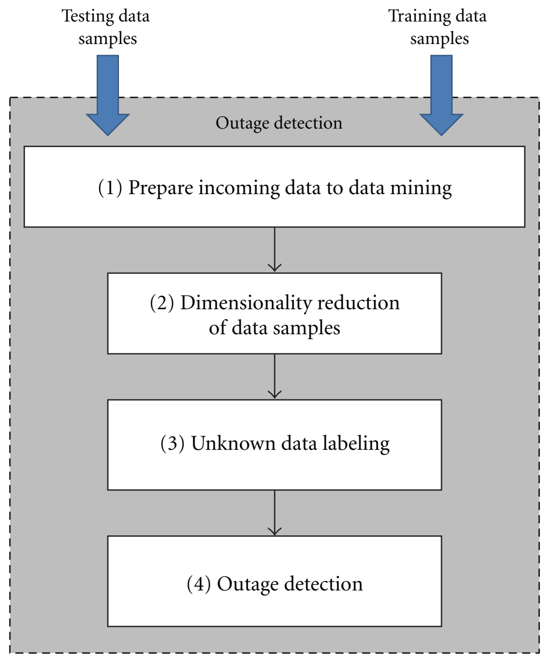


Figure 3: Data-mining outage detection-phase description.

The first step during the learning phase is preprocessing of the arriving MDT measurements which are labeled as periodical, HO-triggered, or RLF-triggered. Labeling is necessary because problem detection in step 4 relies on *supervised learning* from the labeled training samples. Labeling could be done at the eNB before the samples are sent to the TCE. The second step is to check whether or not a proper training database exists. In our case, the requirement is that a sufficient amount of periodical measurements and HO-triggered measurements are gathered from the network during its normal operation. In addition, some RLF samples from previous outage situations are gathered. A training database is created from the preprocessed MDT measurements which characterize normal network behavior without any outages. When the training database is constructed, all new measurement samples are put into the testing database. The operator needs to validate the training database and make sure it really resembles the needed network characteristics for example, the network behavior during its normal operation. The validation could be done by using anomaly detection and unsupervised learning techniques as described in [12].

In the problem-detection phase, recently received MDT measurements in the testing database are compared with the training data to detect anomalous behavior in the system, as depicted in Figure 3. The first step in the outage detection process is to prepare the data in the training and the testing set. Depending on the problem and the applied data mining algorithms, this preprocessing phase may contain several kinds of actions such as data cleaning, data integration, data transformation and data scaling. In our framework, each MDT measurement, as described in Table 1, is cleaned by splitting a single measurement to the header part and the data part. The header part contains information for post-processing of the outage detection results, like visualization and location correlation, but it is not used by the data-mining algorithm. The data part for ith measurement sample is a vector x_i consisting of 10 numerical features as follows:

$$x_i = \{\mathrm{RSRP}_S, \mathrm{RSRP}_{N1}, \ldots, \mathrm{RSRP}_{N3}, \mathrm{RSRQ}_S, \mathrm{RSRQ}_{N1}, \ldots, \mathrm{RSRQ}_{N3}, \mathrm{WCQI}, \mathrm{PHR}\}, \tag{1}$$

where RSRP_S and RSRQ_S are the serving cell measurements and RSRP_{Nj} and RSRQ_{Nj} are the jth strongest neighbor cell measurements, $j = \{1, 2, 3\}$. WCQI is the serving cell wideband CQI measurement and PHR is the serving cell power headroom report. RSRP and RSRQ measurements are given in a logarithmic scale as specified in [14]. Note that in our studies the WCQI and PHR measurements are not exactly the same as in 3GPP specifications. First of all, the CQI represents the downlink wideband signal-to-interference ratio and it is expressed using a dB scale. Moreover, the PHR metric is scaled by the allocation size resulting in a PHR per physical resource block metric as proposed in [15] since it was seen to improve the detection of uplink coverage problems and uplink power control parameterization problems. Thus, the high-dimensional data classifier consisted of 10 features. In addition, the performance of the 10-feature classifier was compared to an 8-feature classifier since the availability of the WCQI and PHR measurements depends on the eNB implementation. The 8-feature classifier uses only UE reported RSRP and RSRQ measurements.

3.2. Dimensionality Reduction. The next step in the outage detection framework is the dimensionality reduction step. The target of the dimensionality reduction is to represent high-dimensional data sets in a lower dimensional space making the data mining faster and less complicated. By having the dimensionality reduction step employed to the framework, the outage detection framework is more robust and can be extended easier with new numerical KPIs. Dimensionality reduction techniques, such as principal component analysis (PCA) are widely used in machine learning.

In our framework, the testing and training data set dimensionality is reduced by using a nonlinear diffusion maps methodology [16–19]. The diffusion maps method allows finding meaningful data patterns in the high-dimensional space and represents them in the lower dimensional space using diffusion coordinates and diffusion distances while preserving local structures in the data. The diffusion coordinates parameterize the high-dimensional data sets, and the diffusion distance provides a local preserving distance metric for the data. In the following, we shortly describe the used dimensionality reduction method originally proposed in [19]:

(i) The data set X is used to construct a unidirectional graph G, where the graph vertices are the data points x and the edges between the data points are defined by a kernel weight function w_ε.

(ii) The diffusion is created by doing a random walk on the graph. The random walk is done from the Markov transition matrix $\mathbf{P}$ which can be obtained by normalizing kernel weight matrix $\mathbf{W}$ with a diagonal matrix $\mathbf{D}$.

(iii) Finally, if $\mathbf{P}$ exists then the Eigen decomposition of the $\mathbf{P}$ can be used to derive the diffusion coordinates $\Psi_t(x_i)$ in the embedded space and the diffusion-distance metric $D_t(x_i, x_j)$.

The kernel weight matrix $\mathbf{W}$ measures the pairwise similarity of the data points in the graph and it must be symmetric, positive, and fast decaying [19]. One common choice for the kernel is:

$$W_{i,j} = w_\varepsilon\left(x_i, x_j\right) = \exp\left(\frac{-\left\|x_i - x_j\right\|}{\varepsilon}\right). \tag{2}$$

If the weight $w_\varepsilon(x_i, x_j)$ between sample x_i and x_j is small it means that points are similar. On the other hand, if the weight is large then the points are different in nature. Variable ε can be used to scale the kernel weight function which on the other hand scales the size of the local neighborhood. In principle, any weight function form of $f(\|x_i - x_j\|)$ fulfilling the above-mentioned criteria could be used to estimate the heat kernel and thus used with the diffusion process [19]. The Gaussian kernel in (2) is scalable and it decays fast, that is, faster than plain Euclidean distance, and therefore it was chosen. Next, the diagonal matrix $\mathbf{D}$ is derived from $\mathbf{W}$ according to

$$D_{i,i} = \sum_{j=1}^{n} w_\varepsilon\left(x_i, x_j\right). \tag{3}$$

If a proper kernel is used, then the matrix $\mathbf{W}$ can be multiplied from left with matrix $\mathbf{D}^{-1}$ to get the normalized Markov transition matrix $\mathbf{P}$:

$$\mathbf{P} = \mathbf{D}^{-1}\mathbf{W}. \tag{4}$$

In the Markov matrix, the $P_{i,j}$, describes the probability to move from sample x_i to sample x_j in the graph with one step. The random walk in the graph is obtained by raising the Markov transition matrix $\mathbf{P}$ to the tth power $\mathbf{P}^t$. This gives the probability to move from sample x_i to sample x_j in the graph with t steps. Finally, the eigen decomposition of $\mathbf{P}^t$ provides tools to define the high-dimensional data set in $\mathbb{R}^n$ in the embedded space $\mathbb{R}^k$ by constructing an estimate of $\mathbf{P}^t$ by using only k largest eigenvectors:

$$\mathbf{P}^t = \sum_{l=1}^{k} \lambda_l^t \psi_l \phi_l^T, \tag{5}$$

where the variables ψ_l and φ_l are right and left eigenvectors, and the variable λ_l is the eigenvalue of the lth eigenvector. Moreover, the diffusion distance D_t^2 and diffusion coordinates Ψ_t can be constructed by using the eigenvalues and the right eigenvectors as proven in [19]:

$$D_t^2\left(x_i, x_j\right) = \sum_{l=1}^{k} \lambda_l^{2t}\left(\psi_l(x_i) - \psi_l\left(x_j\right)\right)^2, \tag{6}$$

where the diffusion distance D_t^2 is the Euclidean distance between the measurement x_i and x_j in the embedded space by using the diffusion coordinates. The diffusion coordinates are constructed using k most significant right eigenvectors and eigenvalues as given in [19]:

$$\mathbf{\Psi}_t(x_i) = \left[\lambda_1^t \psi_1(x_i), \lambda_2^t \psi_2(x_i), \ldots, \lambda_k^t \psi_k(x_i)\right], \tag{7}$$

where the diffusion coordinates $\mathbf{\Psi}_t(x_i)$ for measurement x_i can be obtained from the m-by-k diffusion coordinate matrix $\mathbf{\Psi}_t$. The column vectors of $\mathbf{\Psi}_t$ are the right eigenvectors of $\mathbf{P}^t$ multiplied by the corresponding eigenvalue term λ_l^t as shown in (7). Moreover, the diffusion coordinates for measurement x_i are found in the ith row vector of the diffusion coordinate matrix $\mathbf{\Psi}_t$. As seen from (6) and (7), the diffusion distance for samples in the high-dimensional space corresponds to the Euclidean distance of the samples in the embedded space.

3.3. Data Classification. The third step in our outage detection framework is data classification used to learn the characteristics of the testing data. In our earlier paper, we considered unsupervised learning techniques to detect sleeping cells [12] by incorporating *k means clustering* without taking into account the periodical MDT measurements. In this paper, we are describing the application of the supervised learning classification algorithm known as *nearest neighbors search* (NNS). Difference between the supervised learning and the unsupervised learning techniques is that in the supervised learning we know labeling for the training data and based on the training data characteristics we try to label unknown testing data samples. In our approach the training data consists of samples which belong to one of three class types, labeled as periodical, handover, or RLF samples, and the target is to classify all unknown testing data samples to those three known category types. Motivation of classifying testing data to these three class types is to detect periodical MDT measurements which have similarities with samples belonging to the outage category. By doing the classification, early outage detection can be done even in cases that only insufficient amount of RLF reports are available.

The fundamental idea of NNS is to find a set S_i of nearest neighbors from the training database for each unknown sample x_i in the testing database. One method of determining S_i is to calculate a distance from x_i to all points in the training database. Therefore, the complexity of the NSS depends on the size of training and testing sets as well as the dimensionality of the data samples. In our work, the nearest neighbors search is done in embedded low dimensional space based on the Euclidean distances. This is equally the same as classifying samples in high dimensional space according to the diffusion distances. The set S_i is used to define the labeling for all the unknown samples. There can be a wide range of vendor specific algorithms to determine the label for the unknown samples based on the S_i but here a simple algorithm was used and the class label is chosen based on the largest class in terms of number of samples present in the set S_i.

3.4. Anomaly Detection. The final step of the outage detection framework is anomaly detection. By this stage, the testing database is already labeled and this information is used to detect possible outage or sleeping-cell problems in the network. There are two different principles for detecting anomalous base-station behavior. On one hand, anomalies can be detected in time domain by comparing target base-station behavior in time to the behavior observed earlier. This requires long observation times and data-gathering periods per base station for creating reliable time domain profiles. On the other hand, anomalous base-station behavior can be detected in base station domain by comparing target base-station behavior to the neighboring base stations. In the latter case, more data is gathered in a shorter time period but the data can be biased if the neighboring base stations behave differently, that is, due to the different parameterization. In our framework, the common assumption for all base stations is that during normal operation the amount of RLF samples is small. Thus, the data classification should not result in many periodical MDT samples which are considered to belong to the RLF class. On the other hand, when the network is in outage, many periodical MDT measurements should be similar to the RLF samples. Since the anomaly detection criterion that is, increase of the number of periodical MDT measurements which have similar characteristics as the RLF samples, assumes similar behavior of the base stations during normal operation, the outage detection is based on the base-station domain analysis.

In our framework, the anomaly detection is done by counting the number MDT reports labeled as RLF samples for each eNB and comparing this with the network normal operation in time and base-station domain. The detection is based on the well-known standard score metric which describes how similar an observation of a particular eNB is compared with the normal behavior of a set of neighboring eNBs taking into account the normal deviation of the observations. Standard score z_e for eNB e is defined as,

$$z_e = \frac{|x_e - \mu_x|}{\sigma_x}, \quad (8)$$

where variable x_e is the number of RLF-labeled samples for eNB e and variables μ_x and σ_x are expected mean and standard deviation of the number of RLF-labeled samples in the eNBs local neighborhood. If z_e is much larger than one, then eNB e is probably an anomaly since the amount of RLF-labeled observations do not fit within the normal deviation of the RLF observations.

4. Simulation Results

4.1. Simulation Configuration. Our outage detection approach was verified with the dynamic LTE system simulator which was used to collect a large MDT measurement database. The simulator is capable to simulate E-UTRAN LTE release 8 and beyond in downlink and uplink with several radio resource management, scheduling, mobility, handover, and traffic-modeling functionalities. The simulation scenario consists of a regular hexagonal network layout of 19 sites and 57 base stations with inter site distance of 1750 meters. The 7 center sites are normal cells where the UEs are placed to gather MDT measurements and the outer tier of 12 sites are used only to generate interference. The users were moving in the scenario with velocity of 3 km/h and handover parameters were chosen in a way that the performance during normal operation was assumed to be good. On the other hand, the radio link monitoring values were chosen to trigger the RLF slightly faster than normally to ensure that some RLF samples are gathered during the normal operation of the network. The simulation assumptions are based on the 3GPP macro case 3 specifications [20] defining the used bandwidth, center frequency, network topology, and radio environment as listed in Table 4.

The simulation campaign consisted of a reference and problem simulations. The reference simulation was used to gather training data during the normal operation of the network and the simulated MDT database consisted of 148723 periodical measurement samples, 698 handover samples, and 138 RLF samples. The periodicity of sending MDT measurement reports was 0.5 seconds. In the problem scenario, one eNB was attenuated completely since the target was to model a sleeping cell where the uplink and the downlink are malfunctioning. The outage was created by adding 50 dBi antenna attenuation to the eNB 8. Since all sites were operating on the same band and overlaying interfrequency layer didn't exist, the eNB 8 was in outage. This enlarged the dominance areas of the neighboring cells as depicted in the Figure 4. The dominance area indicates the area where a particular cell is the strongest serving cell. In the left figure, the eNB 8 dominance area is shown with turquoise color, and the size of the area is similar to the other cells. In the right figure, the eNB 8 is sleeping and the area is served by the neighboring cells. Note that the eNB 8 covers less than 5% of the overall area where the UEs are distributed during the simulation.

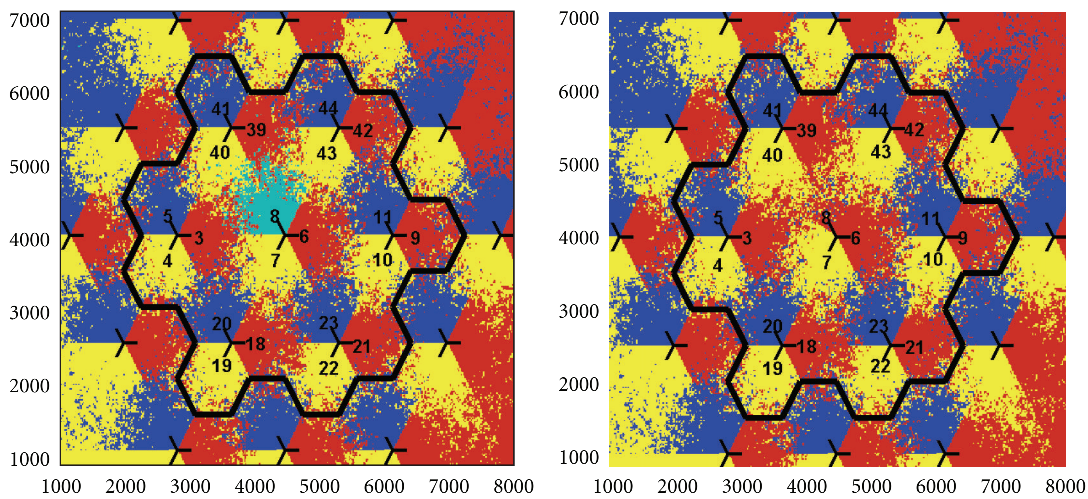
The described dominance area problem is easy to understand, and therefore, it is interesting to see how our approach is able to detect the change in the dominance areas. The MDT database gathered from the reference simulation provides the basis of the training database which defines the statistical structure and the characteristics of three classes. Since the MDT database from the reference simulation was large only a fraction of this data was used in the actual training data set. The training data set was constructed from 3000 periodical samples using random undersampling [21], all HO samples, and all RLF samples. Moreover, the size of the RLF data set was oversampled by a factor of 4 in order to have roughly the same amount of HO and RLF samples in the training data set. Even though oversampling leads to a certain degree of overfitting, and consequently might lead to a degradation of classification accuracy [22], it can also enhance the classifier performance as shown in [11]. All MDT data gathered from the problem simulation is used to construct the unknown database, and each sample in this database is labeled as either periodical, handover, or RLF class as earlier explained in Section 3.

4.2. Simulation Data Mining Results. To be able to detect the anomalous network behavior, all MDT measurement samples in the unknown database was labeled by using the training data set classifier. Labeling of unknown samples was done based on 7 nearest neighbors since this was found to perform reasonably well. The nearest neighbors in the training set were always chosen based on the Euclidean distance in the embedded space which is the same as the diffusion distance in the original space. Classification accuracy of the NNS algorithm applied to MDT data is shown in Tables 2 and 3. The classification accuracy is evaluated with confusion matrices showing the probability of *true-positive* labeling and *false-positive* labeling. Different confusion matrices are shown for 10-feature and 8-feature classifiers. The 10-feature classifier uses all 10 features including WCQI and PHR for the dimensionality reduction as described in (1), whereas the 8-feature classifier uses only UE reported RSRP and RSRQ values. Diagonal cells of the confusion matrices show the true-positive probability indicating the likelihood that a sample is correctly labeled to the same class it belongs. The false-positive likelihood indicates the probability for the samples to be labeled to a wrong class. This kind of comparison is easy to do since we know the real labels of the data. The number inside the parenthesis of the real class column indicates the total amount of the different sample types in simulations and confusion matrices showing how these samples were labeled by the different classifiers.

Table 2 shows the reference simulation labeling accuracy for all MDT samples. It can be seen that the true-positive-labeling likelihood of the reference data is more than 80% for all class types regardless of the used classifier. The 10-feature classifier performs better but the performance of 8-feature classifier is not much worse either. One should note that we do not try to achieve 100% classification accuracy since it is quite likely that some of the periodical, handover, and radio link failure samples would have similar kind of characteristics in any case. Periodical samples are collected in a periodical manner, and therefore, the samples preceding a handover or a radio link failure event are assumed to have similar kind of characteristics. It is worth noting that handovers occur at the cell edge, and depending on the handover parameters and the slow fading conditions some handovers can have similarities with radio link failures.

Classification quality of the MDT samples from the problem and reference simulation is approximately the same as shown in Table 3. Classification accuracy of the 10-feature classifier remains better in the problem scenario as well. There is a small change of 0.8% in handover false-positive labeling but that is negligible since the number of handover samples is only 683 meaning 8 samples were classified differently. A small change in periodical sample false-positive-labeling probability is observed as well. In the problem simulation, the 10-feature classifier labels 0.9% of the periodical samples to radio link failures. This is almost two times higher than in the reference simulation. However, this small difference of 0.4% is significant since the number of periodical samples in the problem simulation MDT database is huge that is, 148693 samples. This means that 1338 additional RLF-like samples were found from the set of periodical MDT samples indicating outage. This is 537% more samples than the 210 true RLFs detected in the problem scenario. If the 8-feature classifier is used the difference is same. However, the classification accuracy is slightly lower, and therefore the total number of RLF-labeled samples is higher in the reference and the problem simulation. On the other hand, the 8-feature classifier can be applied to the interfrequency measurements directly, since it does not use CQI or PHR measurements for the outage detection.

The final goal in the outage detection is to associate RLF-labeled samples with base stations. Generally, MDT samples with detailed location information are reported with latitude and longitude values and rest of the samples can be located based on the RF fingerprint of the MDT measurement. Recall that if only GCI of the serving cell would be used, the detected samples in the dominance area of malfunctioning eNB would be associated with neighboring cells leading to misjudgments. Our assumption is that majority of samples can always be located at least with the accuracy of the dominance area for example, an estimate of the strongest serving cell is known for each sample based on the operators estimate of the dominance areas. In urban network deployments, the definition of dominance areas can become ambiguous due to buildings, street layout, and slow fading. However, since the MDT is used to enhance the network coverage maps, it is assumed that dominance area estimates can be improved in urban environment as well. Therefore, it is



(a) Reference simulation

(b) Problem simulation

Figure 4: Network dominance areas.

Table 2: Reference simulation confusion matrices.

Real class	8 features (only RSRP and RSRQ)			10 features (with CQI and PHR)		
	Per.	HO	RLF	Per.	HO	RLF
Periodical (148723)	96.4%	2.9%	0.7%	96.6%	2.9%	0.5%
Handover (698)	11.9%	82.1%	6.0%	11.8%	83.0%	5.2%
Radio link failure (138)	0.0%	0.0%	100%	0.0%	0.0%	100%

assumed that if the MDT measurements bear the detailed location, the correlation with the dominance areas is not an issue. However, if one of the cells is missing, the positioning and correlation with the RF fingerprint databases could be challenging and even lead to wrong conclusions. In this paper, the inaccurate RF fingerprint positioning is not taken into account, and the results rely on the availability of the MDT reports with detailed location information, that is, latitude and longitude. In Figure 5, the normalized RLF-labeling results from the reference simulation are depicted for all base stations. The RLF-labeled samples are associated with the base stations according to the estimated dominance areas. Blue color refers to periodical samples, green color refers to handover samples, and red color refers to RLF samples which are labeled as radio link failures. The results are normalized with the total number of RLF-labeled samples in the reference scenario. There are a few radio link failures occurring in the reference scenario and only 3% of all RLF-like samples were detected to occur at the dominance area of the eNB 8. These RLFs in the reference scenario are due to the long intersite distances between the base stations and slow-fading effect especially in eNBs 6, 18, and 43.

Based on all RLF-labeled samples, a standard score for each base station is calculated by using (8). The standard score can be used as a simple indicator to detect if eNB behavior is normal or not since it takes into account the statistical variability of the RLF-labeled samples per base station during normal network operation. In Figure 6, standard-score distributions in reference scenario are shown

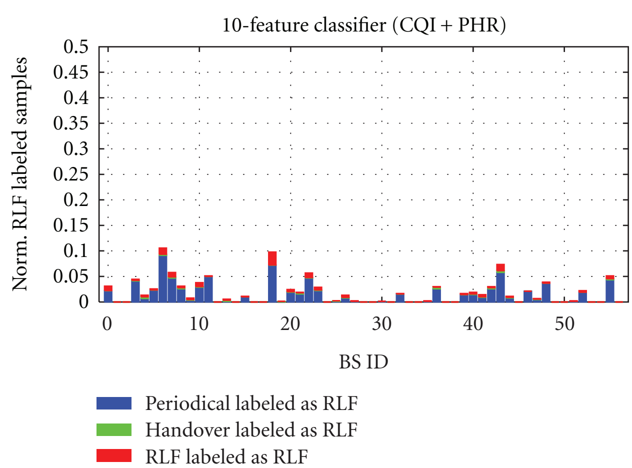


Figure 5: RLF-labeled samples per base station in reference simulation.

for 8-feature classifier with turquoise line and 10-feature classifiers with black-dashed line. Distributions are similar for both classifiers, and 95% of the eNBs have a standard score smaller than two.

The RLF-labeling results in the problem simulation are normalized in a same way as the reference simulation results. After triggering the sleeping-cell problem, the increase in the number of RLF-labeled samples is significant. Figure 7 shows that almost 40% of all RLF-labeled samples were associated

Table 3: Problem-simulation confusion matrices.

Real class	8 features (only RSRP and RSRQ)			10 features (with CQI and PHR)		
	Per.	HO	RLF	Per.	HO	RLF
Periodical (148693)	96.0%	2.9%	1.1%	96.2%	2.9%	0.9%
Handover (683)	13.3%	79.9%	6.8%	12.7%	81.4%	5.9%
Radio link failure (210)	3.3%	0.5%	96.2%	1.0%	1.0%	98.0%

Table 4: Simulator parameters.

Parameter	Notes	Value
3GPP macrocell scenario	Regular cell layout	57 sectors/19 BSs
Intersite distance		1.75 km
Distance-dependent path loss	Macro cell model [20]	$128.1 + 37.6\log_{10}(R_{\text{km}})$
BS Tx power		46 dBm
Slow-fading standard deviation		8 dB
Slow-fading correlation		Site 0.5/Sector 1.0
Fast-fading profile		Typical Urban
UE velocity		3 km/h
UE placement	Uniformly distributed	7 centremost sites
RSRP/RSRQ Measurements	RSRP measurement period	40 ms
	L1 averaging window size	200 ms
	L3 filter coefficient	4
MDT reporting	Periodicity	500 ms
Handover parameters	A3 event threshold	3
	A3 event time to trigger	160 ms
	Handover preparation time	50 ms
Radio link failure monitoring	Qout threshold	−8 dB
	Qin threshold	−6 dB
	T310 timer	600 ms
Number of calls per simulation		4200
Base-stations loading	Full loading in all cells	100% RBs loaded
Diffusion parameter epsilon in (8)	Scales the size of local neighborhood	8
Embedded space dimension	Number of right eigenvectors	6

with the eNB 8 dominance area whereas it was only 3% in the reference scenario. Moreover, it was observed that the total number of RLF-labeled samples is higher for the 8-feature classifier since more periodical samples are labeled as RLFs due to slightly worse classification accuracy. Only 36% of the all RLF-labeled samples were associated with eNB 8 in this case. This indicates that both classifiers detect periodical measurements which are similar with the radio link failures. Moreover, eNB 8 standard score is 26.2 for the 10-feature classifier and 25.2 for the 8-feature classifier. This means that both classifiers detect anomalous network behavior since the standard score is much larger than two-indicating outage. However, the 10-feature classifier is able to isolate the problem from the reference simulation better since the standard score is larger and more RLF-labeled samples were associated with the malfunctioning eNB 8. This indicates that by using CQI and PHR metrics in the classification the outage detection can be improved. On the other hand, the 8-feature classifier can also detect the problem but since it does not depend on the CQI and PHR it can be applied to the interfrequency measurements as well. However, the verification of the interfrequency layer outage is not done in this paper.

Note that UEs in the problem scenario would not detect the presence of the eNB 8. Hence, the existence of the location information and correlation with the dominance information helps to build a better understanding of the root cause and location of the problem. The locations of the RLF-labeled samples in the map grid are illustrated in Figure 8. The simulation area was divided to 42 × 48 meters rectangular map grid points. The number of the RLF-labeled samples were counted for each grid point, and a heat map was used to visualize the likelihoods of the RLF-labeled samples in the estimated dominance area

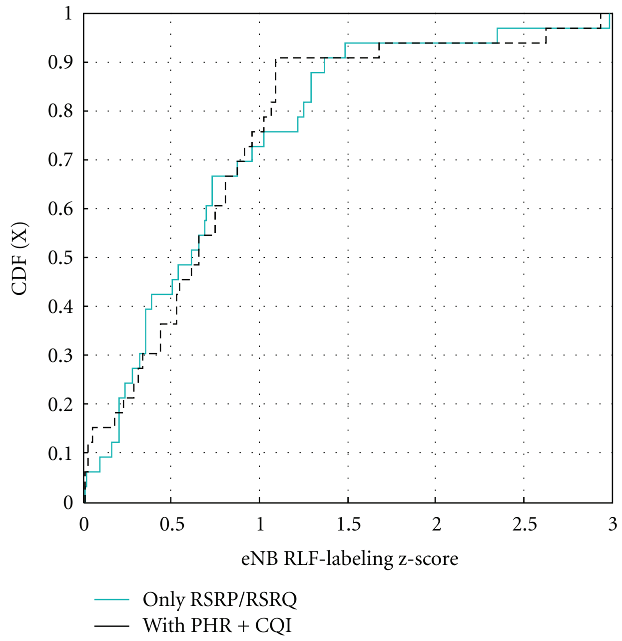


Figure 6: Standard-score distributions for reference simulation.

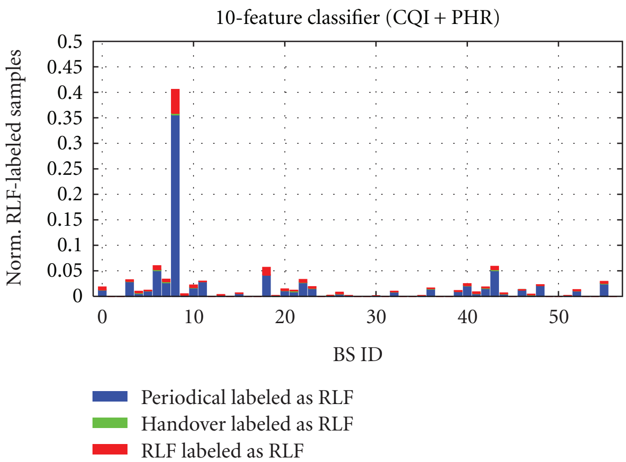


Figure 7: RLF-labeled samples per base station in problem simulation.

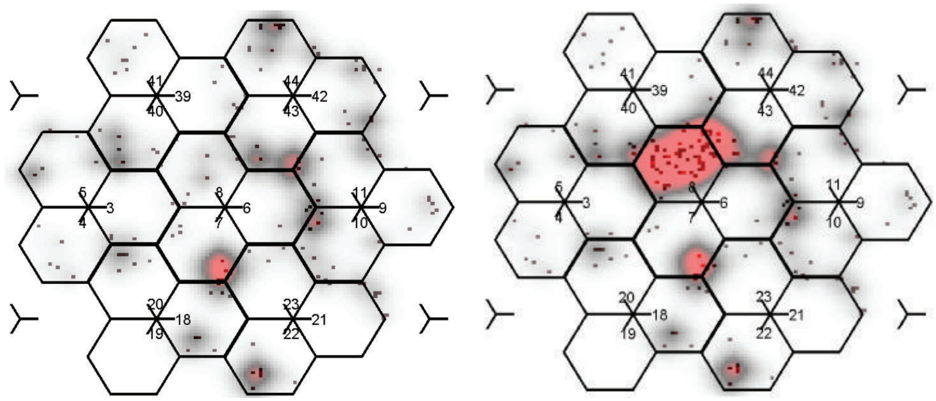
map. A gray color indicates areas in the heat map which might have some outage for example, when approaching a coverage hole, and a bright red color indicates areas where outage is detected. Figure 8(a) shows the heat map for the reference simulation together with estimated dominance areas, and it can be seen that some outage regions at the cell edges do exist due to the slow fading and large ISD between the sites. Figure 8(b) shows the heat map for problem simulation indicating clearly higher likelihood for the outage on the eNB 8 area compared with the reference simulation. It can be seen that the increased likelihood of RLF-labeled samples indicates the change on the dominance areas.

4.3. Anomaly Detection Time. Since anomaly detection is based on the increase in the number of periodical measurements classified as RLF samples, the detection time was analyzed by observing amount of reported samples instead of actual detection time. The amount of reported samples is a better metric, since time needed to gather a sufficient amount of samples for detection depends on the number of active users, user distribution, and MDT configuration, for example, periodicity of the measurements. Average base-station specific z-score metric before and after occurrence of the problem in eNB 8 is depicted in Figure 9.

In Figure 9, the colored curves depict how z-score metric behaves during system simulations in case 10-feature classifier is used. The x-axis indicates the average number of all received MDT reports per eNB, while the simulations advance. The y-axis indicates the eNB z-score as in (8). The z-score values were updated every five seconds but mean and standard deviation values were kept constant according to the reference simulation. Figure 9 indicates that if the observation window is too short, then anomalous base stations are not detected. In the reference simulation before the problem, approximately 3000 MDT samples per eNB are needed until some minor outage is detected. Solid green curve and dotted red curve indicate some outage in eNB 6 and eNB 18. The detection time in this case would depend on the average number of UEs per eNB, their movements in the eNB dominance area, and the periodicity of the MDT reports. For example, if 10 uniformly distributed UEs are sending MDT reports with periodicity of 0.5 second, then the detection for example, reception of 3000 samples, would take 2.5 minutes. The problem triggers after 4500 MDT reports per eNB are received. The eNB z-scores are cleared, and detection is restarted as well. Shortly after triggering the problem, eNB 8 starts to stand out from the statistics. Blue curve indicates that the z-score for eNB 8 is already more than 10 after reception of 1500 MDT reports per eNB. Moreover, purple curve shows that eNB 43 z-score increases from 2 to 3 due to the sleeping cell. This indicates that outage increases slightly in the dominance area of eNB 43 due to the problem in eNB 8. For the eNBs 6 and 18, the outage remains similar.

5. Conclusion

This paper described a data-mining framework which is capable of detecting network outage and sleeping cells in a cellular network by using drive testing databases. The framework is cognitive since it adapts to the deployed network configuration and topology by learning the network characteristics while gathering the training data for the problem classifier. In addition, the described outage detection framework works in a self-organizing manner since it uses the E-UTRAN minimization of drive testing functionality to gather the training and testing databases. The essence of the method is to label unknown data by finding similar characteristics from the previously known network data. For this purpose, diffusion maps dimensionality reduction and nearest neighbors data classification methods were utilized.



(a) Reference simulation (b) Problem simulation

Figure 8: RLF-labeled samples on the map grid.

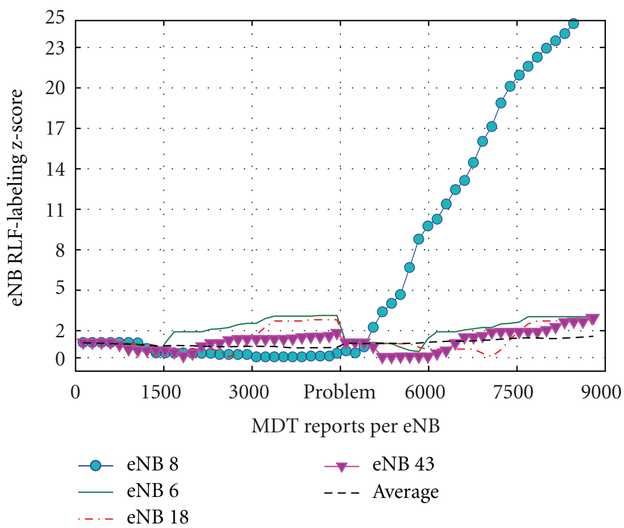


Figure 9: Average eNB z-score before and after problem triggering in eNB 8.

The presented approach is robust since the same principle utilized here can be used for a wide range of different network problems where the problem data can be isolated and used later as known problem classifiers.

In the case of the sleeping cell problem, the detection is based on finding periodical measurements which have similarities with the radio link failures. In the studied verification case, the algorithm gains 537% in the number of samples which can be used for the outage detection in addition to the real radio link failure reports. This makes detection more reliable and possibly faster compared with the algorithms which are based purely on the reported RLF events. Although our approach clearly helps to detect the outage situations by taking into account the periodical samples, there are still some drawbacks in this framework which needs to be solved in the practical deployments. First of all, our approach detects sleeping cells based on the outage present in the dominance areas of the sleeping cell. However, in denser networks, the outage might be less severe and the neighboring base stations can serve the users in the dominance area of sleeping base station without a significant increase of the radio link failures. Moreover, since the typical live networks consist of several overlapping frequency layers, then radio link failures in one layer can be avoided by handing UEs over to another frequency layer. In such situations, the framework could be extended to take into account additional features such as loading level of the cells, the handover activity, or the interfrequency layer measurements. These features together with the change in dominance areas could eventually result in a more comprehensive solution to the sleeping-cell problem. However, one advantage of the presented framework is indeed the robustness due to the dimensionality reduction step. This is a stepping stone for future research allowing an easy inclusion of new features in case of different anomaly detection studies.

Acknowledgments

The authors would like to thank Amir Averbuch and Gil David from Tel Aviv University for their support with the knowledge mining. Moreover, they would like to thank *Renesas Mobile Europe* for use of their system simulator since the work could not have been done without it. In addition, constructive criticism, comments, and support from the colleagues at Magister Solutions Ltd., University of Jyväskylläm and the Radio Network Group at Tampere University of Technology were extremely valuable during the work.

References

[1] 3GPP TS 36.902, "Evolved Universal Terrestrial Radio Access Network (E-UTRAN), Self-Configuring and Self-Optimizing Network (SON) Use Cases and Solutions," v.9.3.1, March 2011.

[2] NGMN Alliance, "Next Generation Mobile Networks Use Cases related to Self Organising Network, Overall Description," v.2.02, December 2008.

[3] 3GPP TR 36.805, "Study on minimization of drive-tests in Next Generation Networks," v.9.0.0, December, 2009.

[4] 3GPP TS 37.320, "Radio measurement collection for Minimization of Drive Tests," v.0.7.0, June 2010.

[5] 3GPP RP-111361, "Enhancement of Minimization of Drive Tests for E-UTRAN and UTRAN—Core Part Approval," Nokia Siemens Networks, Nokia, MediaTek, 2011.

[6] 3GPP TS 32.422, "Subscriber and equipment trace, Trace control and configuration management," v.11.0.1, September 2011.

[7] 3GPP TS 36.331, "Evolved Universal Terrestrial Radio Access (E-UTRA), Radio Resource Control (RRC), Protocol specification," v.10.4.0, December 2011.

[8] M. Amirijoo, L. Jorguseski, T. Kürner et al., "Cell outage management in LTE networks," in *Proceedings of the 6th International Symposium on Wireless Communication Systems (ISWCS '09)*, pp. 600–604, Siena, Italy, September 2009.

[9] M. Amirijoo, L. Jorguseski, R. Litjens, and R. Nascimento, "Effectiveness of cell outage compensation in LTE networks," in *Proceedings of the IEEE Consumer Communications and Networking Conference (CCNC '11)*, pp. 642–647, Las Vegas, Nev, USA, January 2011.

[10] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, 2000.

[11] S. B. Kotsiantis, D. Kanellopoulus, and P. E. Pintelas, "Data Preprocessing for Supervised Learning," *International Journal of Computer Science*, vol. 1, no. 2, pp. 111–117, 2006.

[12] F. Chernogorov, J. Turkka, T. Ristaniemi, and A. Averbuch, *Detection of Sleeping Cells in LTE Networks Using Diffusion Maps*, VTC Spring, Budapest, Hungary, 2011.

[13] C. M. Mueller, M. Kaschub, C. Blankenhorn, and S. Wanke, "A cell outage detection algorithm using neighbor cell list reports," in *Proceedings of the International Workshop on Self-Organizing Systems*, pp. 218–229, 2008.

[14] 3GPP TS 36.214, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer; Measurements," v.10.1.0, March 2011.

[15] J. Turkka and J. Puttonen, "Using LTE power headroom report for coverage optimization," in *Proceedings of 74th IEEE Vehicular Technology Conference (VTC '11-Fall)*, San Francisco, Calif, USA, September 2011.

[16] R. R. Coifman, S. Lafon, A. B. Lee et al., "Geometric diffusions as a tool for harmonic analysis and structure definition of data: diffusion maps," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 21, pp. 7426–7431, 2005.

[17] R. R. Coifman and S. Lafon, "Diffusion maps," *Applied and Computational Harmonic Analysis*, vol. 21, no. 1, pp. 5–30, 2006.

[18] B. Nadler, S. Lafon, and R. R. Coifman, "Diffusion maps, spectral clustering and eigenfunctions of fokker-planck operators," in *Advances Neural Information Processing Systems*, vol. 18, pp. 955–962, 2005.

[19] A. Schclar, "A diffusion framework for dimensionality reduction," in *Soft Computing For Knowledge Discovery and Data Mining*, chapter IV, pp. 315–325, 2008.

[20] 3GPP TS 36.814, "Further advancements for E-UTRA physical layer aspects (Release 9)," v.9.0.0, March 2010.

[21] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 9, pp. 1263–1284, 2009.

[22] R. C. Holte, L. Acker, and B. W. Porter, "Concept learning and the problem of small disjuncts," in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 813–818, 1989.