# A Machine Learning Framework for Detection of Sleeping Cells in LTE Network



Ali Imran
University of Oklahoma


Arsalan Saeed
University of Surrey


Adnan Abu-Dayya
Qatar Mobility Innovations Center

# A Machine Learning Framework for Detection of Sleeping Cells in LTE Network

Ahmed Zoha, Ali Imran, Adnan Abu-Dayya
QMIC
Qatar Science and Technology Park
Doha, Qatar 210531
Email: ahmedz@qmic.com

Arslaan Saeed
CCSR
University of Surrey
Guildford, UK, GU2 9PL

***Abstract*—The rapid advancements in telecommunication systems leads to growing data volume and high customer expectations in terms of cost and quality of service. The changing dynamics of radio network usage poses challenges for the operators in terms of optimizing and maximizing network efficiency while reducing maintenance and operational expenditure. Automatic detection of sleeping cell (SC) (i.e. a cell which is not providing normal services to the users) in the network is one way of lowering maintenance cost and improving network performance. This paper presents an intelligent machine learning framework that make use of minimize drive testing (MDT) functionality to gather key performance indicators (KPI's) of the LTE network. These measurements are further projected to a low-dimensional embedding space and are used in conjunction with state of the art learning models to automate the SC detection process.**

***Keywords*—Sleeping Cell, LTE, Anomaly Detection, Low-Dimensional Embedding***

## I. Introduction

The increased demands of high throughput, coverage and to guarantee quality of service (QoS) incur additional challenges for the network operators. One such challenge is the optimization and maintenance of network performance in a cost-efficient manner. Self-organizing Network (SON) concepts [1] have emerged in the last years, with an aim to introduce intelligent automation in the network. Automation of the network management process through SON concepts as specified in 3GPP Release 10 standards, increases the robustness and efficiency of LTE network, while minimizing the cost of operation. One of the main objective proposed in *Self Healing* functional block of SON is to automate the detection of hardware (HW) or software (SW) failures to enhance the reliability of the system. Conversely, the classical monitoring and fault detection methods are resource consuming and requires human labor. Manual drive testing is required in order to collect radio coverage measurements for monitoring the performance of the network. This is not only time and resource consuming but also have reachability issues as the drive testing is limited to outdoor environments only. To address this limitation while reducing the expenditure of drive testing, minimize drive testing (MDT) functionality [1] is specified in SON by 3GPP.

MDT functionality offers a user equipment (UE) assisted data gathering solution in which a UE is configured to report either periodic or event triggered measurements to the E-UTRAN NodeB (eNB). These measurements includes key performance indicator (KPI's) from the serving and neighboring cells, in addition to time and location information. This user perceived measurements can be leveraged to learn the system behavior and accordingly identify any unexpected deviations in an automated fashion. Upon detection of abnormality in the network behavior, timely compensation actions can be triggered to resolve any issues.

Motivated by this, we propose a machine learning framework to automate the detection of specific network failure, named as Sleeping Cell (SC) in LTE network. SC is a situation when Base Station (eNB) failure is not recognized by the operator as there is no alarm triggered. This situation may occur because of the HW, SW failures at eNB. Failure can further be classified into *logical*(i.e. failure of random access channel procedures) and *hardware* (i.e. break down of eNB component). Such failures are responsible for a cell to become *degraded*, *crippled* or *catatonic*. In this study, we have looked at bidirectional antenna gain failure as a HW fault that causes a cell to become *catatonic*. In a real world scenario, such faults might occur due to the malfunctioning of transmitting and receiving modules at eNB. We have adopted a model-driven approach to automatically identify such situations relying on the MDT measurements forwarded to eNB. In our proposed framework a normal cell behavior is profiled in two stages. Firstly, the UE reported KPI's from a fault-free operating scenario are acquired and further embedded into a lower dimensional subspace. In the next stage, the embedded measurements are then used to train anomaly detection models for identifying abnormal network behavior. In the SC detection stage, the trained models leverage the intrinsic characteristics of embedded representation to finely differentiate between normal and abnormal instances.

Typical method that addresses the problem of SC detection are either based on quantitative models [2] which requires domain expert knowledge, or simply rely on performance deviation metrics for detection [3]. In particular, the problem of detecting catatonic sleeping cells has been addressed using *Neighbor Cell List (NCL)* reports [4]. Until recently, we see an increased interest in applying machine learning methods such as Bayesian Networks [5] and clustering algorithms [6] for detecting cell outages. Although, all the previously mentioned studies tries to address the problem of detecting abnormality in the cell behavior, however our study differs in various aspect. Firstly, we in particularly focus on LTE cellular networks, simulated in accordance to existing 3GPP standards. Secondly, we compare the performance of global and local anomaly detection models for developing a normal cell profile. To the best of our knowledge, we are not aware of any comparative evaluation of such methods for detecting SC in LTE networks.

The organization of the paper is as follows: we discuss our proposed framework for SC detection in Section II. This also includes a brief discussion on two state of the art anomaly detection models namely *k*-nearest neighbor and Local Outlier Factor based Anomaly detector. In Section III, we provide the details of our simulation setup and further report and analyze the results. Finally, we conclude the paper in Section IV.

## II. Our Sleeping Cell Detection Framework

Our three step framework for SC detection in LTE networks has been shown in Figure (1). It consists of acquiring MDT measurements, followed by detection and localization.

1) **Measurement:** In the measurement phase, the training and test datasets are independently collected from the simulation environment. The training dataset consists of KPI's from the normal fault-free operating scenario. The validation and test datasets are acquired from a scenario representing the faulty operation of the network. There are two possible approaches for acquiring the KPI's from the simulation environment as discussed in Section II-A.
2) **Detection** In the detection phase, we first project the training data into a low-dimensional space using Multi-dimensional Scaling (MDS) method [7] as discussed in Section II-B. The MDS embedding of the KPI vectors maximizes the variance in the dataset by increasing the distance between the dissimilar observations and vice versa. Hence, the abnormal KPI's samples lie far from the normal KPI's in the low-dimensional space which results in accurate profiling of normal cell behavior. Therefore, embedded measurements are further used to train global and a local anomaly detection algorithms namely *k*-Nearest Neighbor and Local Outlier Factor, respectively. A brief overview of the algorithms are provided in Section II-C.
3) **Localization** In the localization phase, the location of the sleeping cell is identified based on the classification performed by employed algorithms as further explained in Section II-D.

### A. MDT Measurements for SC Detection

*Minimization of Drive Testing* (MDT) use case for SON were introduced by NGMN during 2008. The idea is to minimize the cost of manual drive testing by enabling the UE to report the coverage measurements. In LTE Release 10 the MDT measurement and reporting schemes have been defined. The MDT measurement functionality allow operator to collect measurements either periodically or event based. A subset of events which generates a MDT report is listed in Table I. The measurement reports consists of cell identification and radio-measurement data as tabulated in Table II. In this study, we have employed A2 and A3 event triggered MDT measurements for SC detection. The simulation parameters are listed in Table III.

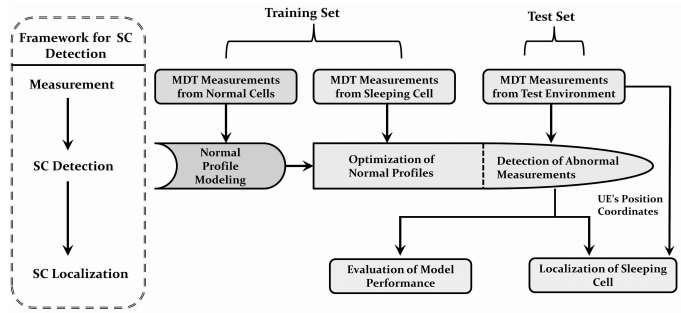


Fig. 1. An three step framework of Sleeping Cell Detection and Localization in LTE Network

| Events | Description |
|---|---|
| A2 | Serving becomes worst than a *threshold* |
| A3 | Neighbor becomes *offset* better than serving |
| RLF | Link quality falls below a certain *threshold* and an interruption in service happens. |

TABLE I. Events employed for SC Detection

### B. Low-Dimensional Embedding

After acquiring MDT measurement reports, the data is further preprocessed by cleaning and scaling it. The last four measurements listed in Table II are combined into single augmented feature vector as shown in Equation 1

$$V = \{RSRP_S, RSRP_{N1}, RSRP_{N2}, ...RSRP_{N3}, \\ ...RSRQ_S, RSRQ_{N1}...RSRQ_{N3}, CQI\} \quad (1)$$

where *S* and *N* stands for serving and neighboring cells, respectively. The 9-dimensional feature vector is further embedded to only two dimensions in the Euclidean space using MDS method. In context of SON, the dimensionality reduction is a crucial step as a high-dimensional KPI database poses challenges for network engineers as well for experts. The real network is complex and dynamic in nature, and it is often not possible to identify few KPIs that really capture the behavior of the system. On the other hand projecting the data onto fewer dimensions of maximum variance uncovers the true structure which ultimately aids the cell profiling process. Moreover, less computational effort is required which consequently leads to low detection delays.

### C. Sleeping Cell Detection

The embedded KPI representation is then used together with state of the art anomaly detection algorithms, which are trained to reject any abnormal test observations which do not conform to the normal network behavior. In one class

| Features | Description |
|---|---|
| Time and Location | Time stamp and longitude and latitude information |
| Serving Cell info | Cell Global Identification (CGI) |
| RSRP | Reference Signal Received Power in dBm |
| RSRQ | Reference Signal Received Quality in dB |
| Neighboring Cell Information | Three Strongest intra-LTE RSRP, RSRQ information |
| CQI | Serving Cell Channel Quality Indicator |

TABLE II. Structure of MDT measurements

classification framework, normal observations are used to train the anomaly detection models so that test instances can be classified as either belonging to a normal class or vice versa by computing a threshold '$\theta$' based on a certain dissimilarity measure '$\mathcal{D}$' between the two:

$$f(x_i) = \begin{cases} Normal, & \text{if } \mathcal{D}(x_i, D_{train}) \leq \theta \\ Abnormal, & \text{if } \mathcal{D}(x_i, D_{train}) \geq \theta \end{cases} \tag{2}$$

where $D_{train}$ is a subset of dataset $D$. The dissimilarity criteria used by $k$-Nearest Neighbor and Local Outlier Factor based models are briefly summarized as follows:

*1) k-Nearest Neighbor based Anomaly Detector (k-NNAD):* Let $x_i$ be the test instance, and $k$ be the $k^{th}$ neighbor in the training set $D_{train}$. To label $x_i$ as normal or abnormal, the *KNNAD* computes a $\mathcal{D}_{\mathcal{KNNAD}}$ based on Equation 3

$$\mathcal{D}_{\mathcal{KNNAD}}(x_i, k, D) = \frac{1}{N_{tr}} \sum_{i=1}^{N_{tr}} \mathcal{I}(d_t \leq d_i) \tag{3}$$

The $N_{tr} = \mid D_{train} \mid$, and $d_t$ is the distance of $x_i$ from its $k^{th}$ nearest neighbor in the training set $D$, whereas $d_i$ is the distance between $i$ and its $k^{th}$ nearest training object in $D_{train}$. Equation 3 represents a global anomaly detection score as proposed in [8], which is compared against $\theta$ to mark the test instance as anomalous or otherwise.

*2) Local Outlier Factor based Anomaly Detector (LOFAD):* The LOFAD [9] tries to compare the local density $\rho$ of the object to that of its $k$ neighbors. It constructs a local neighborhood of an instance $x_i$ and defines its distance to $k^{th}$ nearest neighbor $NN(x_i, k)$:

$$d_b(x_i, k) = d(x_i, NN(x_i, k)) \tag{4}$$

The $d_b(x_i, k)$ is used to construct a neighborhood $\mathcal{N}(x_i, k)$ by including all those points in the neighborhood fulfilling the following criteria: $d(x_i, x_j) \leq d_b(x_i, k)$. Formally, reachability distance $d_r$ is defined to estimate the $\rho(x_i, k)$ as follows:

$$d_r(x_i, k) = \max\{d_b(x_j, k), d(x_j, x_i)\} \tag{5}$$

and $\rho$ can be defined as

$$\rho(x_i, k) = \frac{\mid \mathcal{N}(x_i, k) \mid}{\sum_{x_j \in \mathcal{N}(x_i,k)} d_r(x_i, x_j, k)} \tag{6}$$

The $d_r(x_i, x_j, k)$ ensures that instances that lie farther away from $x_i$ have lesser impact on $\rho(x_i, k)$. Finally the $\mathcal{D}$ can be calculated by comparing the $\rho$ of $x_i$ to its $\mathcal{N}(x_i, k)$, formally defined as:

$$\mathcal{D}_{\mathcal{LOFAD}}(x_i, k, \mathcal{D}_{train}) = \frac{\sum_{x_j \in \mathcal{N}(x_i,k)} \frac{\rho(x_j,k)}{\rho(x_i,k)}}{\mid \mathcal{N}(x_i, k) \mid} \tag{7}$$

The $\mathcal{D}_{\mathcal{LOFAD}}$ will be 1 if $x_i$ lie inside a cluster or else it receives a higher value which can be compared against $\theta$ to label it as an anomaly.

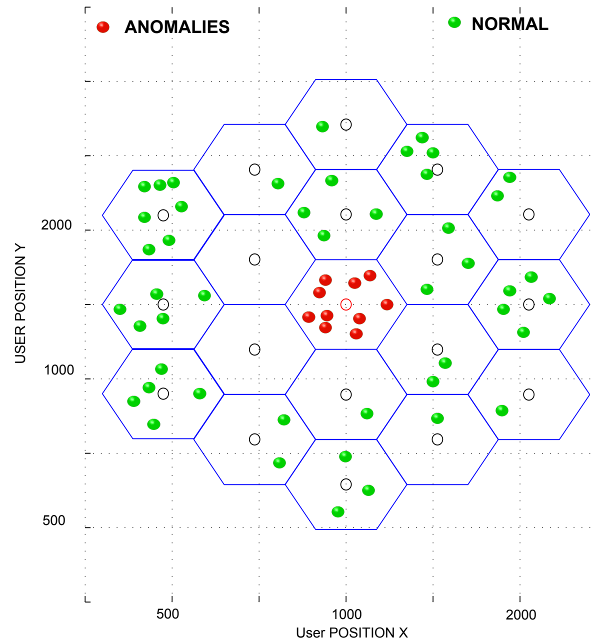


Fig. 2. SC localization based on UE reported position information

## D. Localization of SC

After the employed detection algorithms classify the test measurements, the UE reported position information is utilized to perform localization of SC. As shown in Table II, the MDT measurement reports also contains time and location information. These two are not used at the SC detection stage as shown in Equation 1. However, based on the coordinate information, the classified measurements can be further mapped to network topology. As a result, the cell which corresponds to the highest number of abnormal measurements can be easily identified as shown in Figure 2.

# III. Performance Evaluation

## A. Simulation Setup

A full dynamic system tool is employed to simulate the LTE network based on 3GPP specifications. Two reference scenarios are generated to collected normal and abnormal measurements. In a SC scenario, the antenna gain of a problematic cell is degraded to $-50$ dBi as indicated in Table III.

## B. Analysis of Detection Performance

To train and evaluate the accuracy of our detection models, as pointed out earlier the training observations are acquired from a normal operating scenario. However, the test dataset is divided into (30%) validation and (70%) test sets. We combine the training and validation sets and has applied 10-fold cross-validation to optimize the parameters (i.e. $k = 1, 2, ...10$) for $k$-NNAD and LOFAD, ($\alpha = 0.1, ...1.0$). A standard evaluation method, namely Area Under the Curve (AUC) measure associated with Receiver receiver operating characteristic (ROC) curve [10] is adopted to evaluate the performance of our target

| Parameter | Values |
| --- | --- |
| Cellular Layout | Macro World 27 sites |
| Sectors | 3 |
| User Distribution | Uniform |
| Path Loss | $L[dB] = 128.1 + 37.6log_{10}(R)$ |
| Antenna Gain (Normal Scenario) | 15 dBi |
| Antenna Gain (SC Scenario) | $-50$ dBi |
| Slow Fading Std | 8 dB |
| Simulation Length | 100s (1 time step = 71.43$\mu$s) |
| BS Tx Power | 46 dBm |
| Network Synchronization | Asynchronous |
| HARQ | Asynchronous, 8 SAW channels, Maximum Retransmission = 3 |
| Cell Selection Criteria | Strongest RSRP defines the target cell |
| Load | 20 users/cell |
| Traffic Model | Infinite Buffer |
| HO Margin | 3dB |
| HO trigger time | 256 ms |

TABLE III. SIMULATION PARAMETERS

| Model | Approach | AUC score |
| --- | --- | --- |
| $k$-NNAD | Global | 85± 2.3 |
| LOFAD | Local | 76 ± 3.8 |

TABLE IV. PERFORMANCE OF TARGET ANOMALY DETECTION MODELS FOR SC DETECTION

models. We report the performance of each model in Table IV, which is based on optimal parameter setting of the algorithms. The main difference between the employed algorithms are their method to compute the $\mathcal{D}$, which determines their approach as local or global. Local anomalies are localized to a small spatial region (i.e. local density) or a neighborhood whereas global anomalies are bounded by entire dataset (i.e. global densities). It has been observed that the KPI's from the normal scenario when projected to an Euclidean space are grouped into dense neighborhood, whereas the measurements obtain from the abnormal or sleeping cell scenario lie far from this neighborhood. This is because MDS tries to maximize the variance between the data points and dissimilar points are projected far from each other. Thus, the abnormal measurements acts as global anomalies. This explain the reason that $k$-NNAD being a global anomaly detector has outperformed LOFAD, and achieve high detection accuracy. Moreover, we found out that an average difference of 35 dBm is observed between the RSRP values of a normal and a sleeping cell.

The density value of abnormal measurements is low, however some of them do overlap with micro clusters of normal measurements. These micro clusters lie at the border of dense neighborhood and LOFAD wrongly treat them as local outliers, and thus achieve low detection scores. The results clearly indicate that global models are well suited for detecting abnormal behavior of the network in the embedded space in comparison to local models. MDS embedding of the KPI measurements offers a clear advantage, as regular occurring observations form a clear grouping in the lower-dimensional space. This aids the anomaly detection algorithms in isolating normal network measurements from the abnormal ones with high accuracy. Once, we automatically classify the measurements as normal or abnormal , they can further be mapped to network to localize the position of sleeping cell as discussed in Section II-D.

## IV. CONCLUSION AND FUTURE WORK

In this paper, we present a machine learning framework for automating the sleeping cell detection process in an LTE network. Our proposed approach first acquire key performance measurements from the fault-free operating network. The data is further embedded into a lower-dimensional space. The embedded measurements are used to build a normal profile of the network by training the $k$-NNAD and LOFAD detection models. The models are later used to automatically detect abnormal measurements from the test scenario. The detection accuracy of $k$-NNAD is found to be much higher than LOFAD due to its global detection approach. Finally the UE reported coordinate information is employed to localize the position of sleeping cell. This is a preliminary study and in future we aim to extend our research for heterogeneous networks.

## ACKNOWLEDGMENT

This work was made possible by NPRP grant No. 5-1047-2437 from the Qatar National Research Fund (a member of The Qatar Foundation). The statements made herein are solely the responsibility of the authors.

## REFERENCES

[1] S. Hämäläinen, H. Sanneck, C. Sartori *et al.*, *LTE Self-Organising Networks (SON): Network Management Automation for Operational Efficiency*. John Wiley & Sons, 2012.

[2] 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.

[3] B. Cheung, S. G. Fishkin, G. N. Kumar, and S. A. Rao, "Method of monitoring wireless network performance," Sep. 21 2004, uS Patent App. 10/946,255.

[4] C. M. Mueller, M. Kaschub, C. Blankenhorn, and S. Wanke, "A cell outage detection algorithm using neighbor cell list reports," in *Self-Organizing Systems*. Springer, 2008, pp. 218–229.

[5] 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," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 4, pp. 2451–2461, 2008.

[6] Y. Ma, M. Peng, W. Xue, and X. Ji, "A dynamic affinity propagation clustering algorithm for cell outage detection in self-healing networks," in *Proceedings of IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2013, pp. 2266–2270.

[7] T. F. Cox and M. A. Cox, *Multidimensional scaling*. CRC Press, 2010.

[8] M. Zhao and V. Saligrama, "Outlier detection via localized p-value estimation," in *In Proceedings of 47th Annual Allerton Conference on Communication, Control, and Computing, 2009*. IEEE, 2009, pp. 1482–1489.

[9] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, "Lof: identifying density-based local outliers," in *ACM Sigmod Record*, vol. 29, no. 2. ACM, 2000, pp. 93–104.

[10] A. P. Bradley, "The use of the area under the roc curve in the evaluation of machine learning algorithms," *Pattern recognition*, vol. 30, no. 7, pp. 1145–1159, 1997.