DCAD: Dynamic Cell Anomaly Detection for Operational Cellular Networks

Gabriela Ciocarlie, Ulf Lindqvist, Kenneth Nitz
SRI International
Menlo Park, California, USA
{gabriela.ciocarlie,ulf.lindqvist,kenneth.nitz}@sri.com

Szabolcs Nováczki
NSN Research
Budapest, Hungary
szabolcs.novaczki@nsn.com

Henning Sanneck
NSN Research
Munich, Germany
henning.sanneck@nsn.com

***Abstract*—The Self-Organizing Networks (SON) concept includes the functional area known as self-healing, which aims to automate the detection and diagnosis of, and recovery from, network degradations and outages. In this paper, we present Dynamic Cell Anomaly Detection (DCAD), a tool that implements an adaptive ensemble method for modeling cell behavior [5], [6]. DCAD uses Key Performance Indicators (KPIs) from real cellular networks to determine cell-performance status; enables KPI data exploration; visualizes anomalies; reduces the time required for successful detection of anomalies; and accepts user input.**

***Index Terms*—Self-Organizing Networks (SON), cell anomaly detection, Self-Healing, performance management, Key Performance Indicators (KPIs)**

I. Introduction

Self-Organizing Networks (SON) [1] provide increased automation of network operations with optimized resource utilization to meet cellular network users' expectations for "unlimited" capacity and "ubiquitous" coverage. Among different components, the SON architecture includes configuration, optimization, and troubleshooting capabilities that aim to satisfy self-configuration, self-optimization, and self-healing requirements.

In this paper, we focus on demonstrating self-healing capabilities, which reduce the operator effort and the outage time to provide faster maintenance. Dynamic Cell Anomaly Detection (DCAD) builds on and demonstrates previous work [5], [6] in analyzing Key Performance Indicators (KPIs), which are highly dynamic measurements of cell performance, to determine the state of a cell. DCAD combines different classifiers and classifies new data points by taking a weighted vote of their prediction. Moreover, it enables KPI data exploration; visualizes anomalies; reduces the time required for successful detection of anomalies; and accepts user input.

II. Cell Anomaly Detection

Our cell anomaly detection framework [5], [6] aims to determine the relevant features required for detection of anomalies in cell behavior based on the KPI measurements. The main hypothesis is that no single traditional time-series anomaly detection method (classifier) can provide the desired detection performance. This is due to the wide range in both the types of KPIs that need to be monitored and the network incidents that need to be detected. Consequently, we proposed an ensemble method, which combines different classifiers [5], [6].

978-1-4799-0913-1/14/$31.00 © 2014 IEEE

The ensemble-method framework applies individual univariate (applied to each KPI) and multivariate (applied to multiple KPIs) methods to the training KPI data, leading to the construction of a pool of different predictors. We use different univariate methods for modeling the KPI behavior and to test data against the built models, which use Empirical Cumulative Distribution Functions (ECDF) [2], Support Vector Machines (SVM) [3] and Autoregressive, Integrated Moving Average (ARIMA) models. Our framework uses SVM and Vector Auto-Regressive (VAR) models for the multivariate case [4].

Using the pool of predictors, the predictions obtained on the KPI data "under test" (i.e., subject to detection) along with the weights allocated to each predictor lead to the computation of the KPI degradation level (i.e., the deviation of a KPI from its "normal" state). The proposed methods rely on context information (available for cellular networks) extracted from human-expert knowledge, Configuration Management (CM) data, or confirmed Fault Management (FM) input data to make informed decisions. We define confirmed FM data as the machine-generated alarms that were confirmed by human operators.

III. DCAD Tool

DCAD relies on a main dashboard (Figure 1), which enables the investigation of the detection performance of each individual classifier and of the ensemble method at the cell and KPI levels. DCAD also provides means to compare methods, respectively KPIs, against each other. These exploration capabilities become instrumental in providing insight information of the KPIs characteristics and providing the human operator with relevant information on cell status.

DCAD allows the human operator to:

- visualize the performance of each individual univariate and multivariate method, and the ensemble method, illustrating both the raw KPI measurements and the result of each method as a numerical measure referred to as the KPI degradation level, which indicates the severity of the degradation;
- investigate different KPIs and cells in the network and provide geographical information for the cell in scope, illustrated on a map;

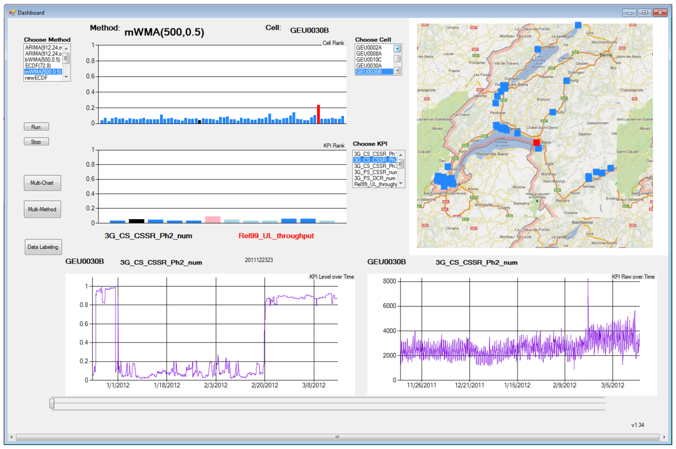

Fig. 1. DCAD main dashboard, which allows selection of the detection methods and investigation of their KPI degradation levels; illustration of cells' geographical location; and comparison of different characteristics of cells and methods.

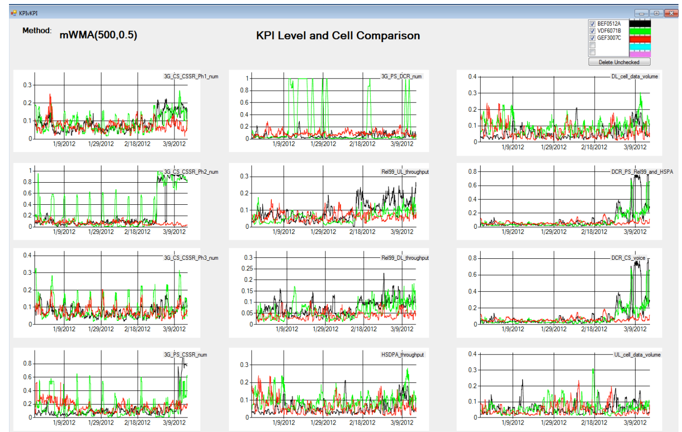

Fig. 2. Cell comparison window comparing all KPI degradation levels for as many as five cells selected in the main dashboard.

Fig. 3. Data labeling (top) and KPI degradation level computation (bottom). In red, the precomputed KPI degradation levels; in blue, the KPI degradation levels computed based on user input (marked as blue in the top graph).

- illustrate the overall state of each cell across all the KPIs to indicate the worst-performing KPIs;
- compare the KPI degradation levels belonging to selected cells, for a more detailed investigation of cell status (Figure 2);
- compare different methods applied to different cells, for a more detailed investigation of the proposed methods;
- provide *ground truth* input, defined as labels associated with the data points that indicate whether or not the data represents a real problem, based on expert knowledge;
- and provide labels to the ensemble method, e.g., some external conditions like a special event causing high traffic load. The top graph in Figure 3 illustrates the raw KPI measurement for a particular cell, which after 2/28/2012 exhibits a significant drop in value. When labeled data (illustrated in blue after the value drop) is available, our framework uses it to adapt and make better predictions. The bottom graph illustrates the KPI degradation level computed with and without considering the labeled data. We observe that, without the labels, the system would deem the data after 02/28/2012 as abnormal (given the high KPI degradation value in red); when using labels, our system adapts to the change (given the low KPI degradation level in blue). When labels are available, the ensemble method creates new models for the pool and uses the label information to adjust the weights accordingly (since it is a dynamic method).

Our demonstration will capture all these features applied to real cellular network data and providing situational awareness as part of an automated mechanism for taking recovery actions.

ACKNOWLEDGMENT

We thank Lauri Oksanen, Kari Aaltonen, Richard Fehlmann, Christoph Frenzel, Péter Szilágyi, Michael Freed, and Christopher Connolly for their contributions.

REFERENCES

[1] S. Hämäläinen, H. Sanneck, and C. Sartori (eds.), "LTE Self-Organizing Networks (SON): Network Management Automation for Operational Efficiency," Wiley, 2012.
[2] S. Nováczki, "An Improved Anomaly Detection and Diagnosis Framework for Mobile Network Operators," Conference on Design of Reliable Communication Networks (DRCN 2013), Budapest, Mar. 2013.
[3] S. Rüping, "SVM Kernels for Time Series Analysis," In LLWA 01—Tagungsband der GI-Workshop-Woche Lernen—Lehren—Wissen—Adaptivität, Forschungsberichte des Fachbereichs Informatik der Universität Dortmund, Dortmund, Germany, 2001.
[4] B. Pfaff, "VAR, SVAR and SVEC Models: Implementation Within R Package vars," Journal of Statistical Software, Vol. 27, Issue 4, 2008.
[5] G. Ciocarlie, U. Lindqvist, S. Nováczki, H. Sanneck, "Detecting Anomalies in Cellular Networks Using an Ensemble Method," 9th International Conference on Network and Service Management (CNSM), Zürich, Switzerland, 14–18 Oct. 2013, pp. 171–174.
[6] G. Ciocarlie, U. Lindqvist, Kenneth Nitz, S. Nováczki, H. Sanneck, "On the Feasibility of Deploying Cell Anomaly Detection in Operational Cellular Networks," To appear in IEEE/IFIP Network Operations and Management Symposium (NOMS), Krakow, Poland, May 2014.