Detecting Anomalies in Cellular Networks Using an Ensemble Method

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

Szabolcs Nováczki
Nokia Siemens Networks Research
Budapest, Hungary
szabolcs.novaczki@nsn.com

Henning Sanneck
Nokia Siemens Networks 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. This paper focuses on the problem of cell anomaly detection, addressing partial and complete degradations in cell-service performance, and it proposes an adaptive ensemble method framework for modeling cell behavior. The framework uses Key Performance Indicators (KPIs) to determine cell-performance status and is able to cope with legitimate system changes (i.e., concept drift). The results, generated using real cellular network data, suggest that the proposed ensemble method automatically and significantly improves the detection quality over univariate and multivariate methods, while using intrinsic system knowledge to enhance performance.**

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

I. Introduction

The need for adaptive, self-organizing heterogeneous networks is particularly apparent given the explosion of mobile data traffic (Chapter 10 in [1]) that stems from increased use of smartphones, tablets, and netbooks for day-to-day tasks. The expectations for mobile networks have grown along with their popularity, and include ease of use, high-speed data transmission, and responsiveness. Heterogeneous Networks (HetNet) combining different Radio Access Technologies (RATs) (3G, LTE, WiFi) and different cell layers (macro, micro, pico) within those RATs can offer these capabilities, providing virtually unlimited capacity and ubiquitous coverage. However, a high degree of distribution introduces a high level of complexity requiring additional mechanisms, such as Self-Organizing Networks [1], to manage that complexity.

A. Self-Healing for SON

This paper focuses on self-healing capabilities, which reduce operator effort and outage time, thereby providing faster maintenance. Specifically, the problem that we address is automatic cell anomaly detection. Typically, research has focused only on Cell-Outage Detection (COD) [2] and Cell-Outage Compensation (COC) [3] concepts, but, more recently, detection of general anomalies has also been addressed [4]. This paper addresses both the outage case and the case where the cell can provide a certain level of service, but its performance has degraded to a point below an expected tolerable level and directly impacts users' experience.

B. Contributions

The key challenge for addressing the more general problem of cell degradation is creating a robust method for modeling normal cell behavior. This approach uses Key Performance Indicators (KPIs), which are highly dynamic measurements of cell performance, to determine the state of a cell. KPIs require modeling techniques that can cope with concept drift, defined as the phenomenon where the normal behavior of the system legitimately changes over time (e.g., by the increasing amount of user-induced traffic demand).

This paper proposes a novel method for modeling cell behavior to help address these problems. Our implementation and experiments focus on the problem of creating adaptive models, leveraging the intrinsic characteristics of the environment where the models are created. The work described here provides several contributions by:

- proposing a new ensemble-method approach for cell anomaly detection that computes a numerical measure referred to as the KPI degradation level [5], to indicate the severity of the degradation,
- using intrinsic knowledge of the system to enhance the ensemble-method learning in order to cope with concept drift and provide automation,
- building a system to implement the algorithms, applying the system to a real KPI dataset, and analyzing the performance of the proposed framework.

II. Cell Anomaly Detection

The first goal of the proposed framework is determining the relevant features needed for detecting anomalies in cell behavior based on the KPI measurements. Because KPIs are measurements that are collected as ordered sequences of values of a variable at equally spaced time intervals, they constitute a time series and can be analyzed with known methods for time-series analysis. An anomaly in a time series can be either a single observation or a subsequence of a time series with respect to a normal time series. *Testing* is defined as the comparison of a set of KPI data to a model of the normal state established by an earlier observed set of KPI data referred to as *training* data. *Ground truth* is defined as the labels associated with the data points that indicate whether or not the data represents a real problem.

ISBN 978-3-901882-53-1, 9th CNSM and Workshops ©2013 IFIP

Our hypothesis is that no single traditional time-series anomaly detection method (classifier) could provide the desired detection performance. This is due to the wide range in the types of KPIs that need to be monitored, and the wide range of network incidents that need to be detected.

The proposed ensemble method combines different classifiers and classifies new data points by taking a weighted vote of their prediction. It effectively creates a new compound detection method that, with optimized weight parameter values learned by modeling the monitored data, can perform significantly better than any single method.

A. *Univariate Time-Series Analysis*

Individual KPIs collected for each cell are univariate time series that can be analyzed with the following methods:

- Using a sliding window, an Empirical Cumulative Distribution Function (ECDF) [6] is computed for each window. In the training phase, sliding windows that are similar based on the Kolmogorov-Smirnov (KS) test are captured in clusters represented by a centroid. In the testing phase, each sliding window is tested against the centroids of the clusters and KPI degradation level is defined as the minimum distance from the centroids.
- A Support Vector Machine (SVM) [7] method is used to build KPI models. The training windows are used to build one-class SVMs [8] with a radial basis function (RBF) kernel. In the testing phase, the anomaly score of a test window is 0 or 1, depending on whether it is classified as normal (score of 0) or anomalous (score of 1). The KPI degradation level is computed as the normalized value of abnormal sequences in a number of consecutive tests.
- Using a predictive approach, KPI behavior is captured by autoregressive, integrated moving average (ARIMA) models. Seasonal components are removed using STL, a Seasonal-Trend decomposition procedure based on Loess [9]. STL is robust to outliers, meaning that noise will not affect the seasonal and the trend components, but only the residual component. Two different implementations of the ARIMA modeling are used: static "o," in which only one model is created; and dynamic "m," in which multiple models are created over time.

B. *Multivariate Time-Series Analysis*

The set of all KPIs collected for each cell is considered a multivariate time series that can be analyzed with the following methods:

- Using a sliding window, multivariate one-class SVM models are built across all time series. In the testing phase, their output is just a label with the value *normal* or *abnormal*. This approach provides a high-level view of the KPIs' behavior as a whole without providing a severity indication for each KPI. This multivariate method is relevant for the ensemble-method framework, in which the multivariate prediction is considered when generating individual KPI degradation levels.
- Using a predictive approach, Vector Auto-regressive (VAR) models are applied for the multivariate case. VAR is a statistical model that generalizes the univariate AR model [10]. The VAR approach generates a model for each KPI of a cell while capturing the linear interdependencies among all KPIs (i.e., each KPI is expressed in relationship to all the other KPIs). The VAR models enable seasonal adjustment. Two different implementations of the VAR modeling are used: static "o," in which only one model is created; and dynamic "m," in which multiple models are created over time.

The computation of KPI degradation levels for both multivariate SVM and VAR models is analogous to the case of univariate ARIMA and SVM models.

C. *Ensemble Method for Cell Anomaly Detection*

The proposed ensemble-method framework applies individual univariate and multivariate methods to the training KPI data and relies on context information (available for cellular networks) extracted from human-generated Configuration Management (CM) or confirmed Fault Management (FM) input data to make informed decisions. Confirmed FM data is defined as the machine-generated alarms that were confirmed by human operators.

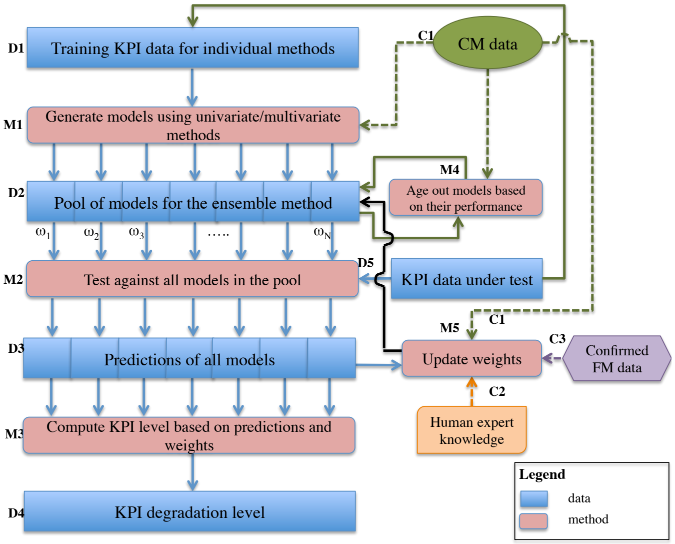
D1
Training KPI data for individual methods
C1
CM data
M1
Generate models using univariate/multivariate methods
D2
Pool of models for the ensemble method
M4
Age out models based on their performance
ω_1
ω_2
ω_3
.....
ω_N
D5
M2
Test against all models in the pool
KPI data under test
M5
C1
D3
Predictions of all models
Update weights
C3
Confirmed FM data
C2
M3
Compute KPI level based on predictions and weights
Human expert knowledge
D4
KPI degradation level
Legend
data
method

Fig. 1. Overall approach of the proposed ensemble method applied to a single cell in a cellular network. Data is depicted in blue rectangles and methods in pink rectangles with rounded corners. The remaining elements indicate different context information. The dashed lines indicated that an event is triggered in the presence of new evidence/data

Figure 1 presents the details of the proposed ensemble method, which implements a modified version of the weighted majority algorithm (WMA) [11]. The modified WMA returns a KPI degradation level in the range [0,1] and uses context information for updating the weights and creating new models.

- Initially, for a given time period, the KPI measurements of a given cell are selected as the training dataset (**D1**) for the pool of models of the ensemble method.

- A diverse set of univariate and multivariate algorithms (**M1**) is applied to the training dataset (**D1**).
- The result of (**M1**) is a set of models used as the pool of models for the ensemble method (**D2**). Each model in the pool of models has a weight, ω_i, associated with it. For the initial pool of models, all models have the same weight value assigned ($\omega_i = 1$).
- Given the pool of models (**D2**), the stream of KPIs is used in a continuous fashion as the testing dataset (**D5**). Any CM change (**C1**) triggers the testing dataset to also become the training KPI dataset, after which the method for generating a new set of models (**M1**) is executed.

  If the pool of models reaches the maximum number of models, the CM change also triggers an exponential decay aging mechanism (**M4**), which removes models from the pool based on both their age and performance (according to $\omega_i * \alpha^{age_i}$, where $0 < \alpha < 1$ and age_i is the number of hours since the model was created).
- The testing dataset (**D5**) is tested against the models in the pool of models using the testing techniques corresponding to the univariate and multivariate methods (**M2**).
- The result of (**M2**) is a set of KPI-degradation-level predictions provided by each individual model in the pool of models (**D3**).

  Ground truth information updates (human-expert knowledge (**C2**), confirmed FM data (**C3**), and CM change information (**C1**)) trigger the update weights method (**M5**), which penalizes the models in the pool of predictors based on their prediction with regards to the ground truth ($\omega_i \leftarrow \beta{*}\omega_i$, where $\beta \in [0, 1]$). The human-expert knowledge assumes a manual process; while the confirmed FM data usage and the CM change detection are automated processes. The result of (**M5**) is an updated pool of models (**D2**) with adjusted weights, which continue to be used in the testing mode.
- All the predictions in (**D3**) along with the weights associated with the corresponding models are used in a modified weighed majority approach (**M3**) to generate the KPI degradation level, where $\tau \in [0, 1]$ is the threshold that determines whether data is deemed normal or abnormal.

$$q_0 = \sum_{KPI<\tau} \omega_i,\; q_1 = \sum_{KPI\geq\tau} \omega_i$$

- The result of (**M3**) is the KPI degradation level (**D4**) associated with each KPI measurement of each cell.

$$\overline{KPI\_level} = \begin{cases} \dfrac{\sum_{KPI\geq\tau} \omega_i{*}KPI\_level_i}{\sum_{KPI\geq\tau} \omega_i}, \text{if } q_1 > q_0 \\ \\ \dfrac{\sum_{KPI<\tau} \omega_i{*}KPI\_level_i}{\sum_{KPI<\tau} \omega_i}, \text{if } q_1 \leq q_0 \end{cases} \qquad (1)$$

III. Evaluation of Ensemble Method

This section quantifies the increase in detection accuracy when the ensemble method is applied to the proposed univariate and multivariate methods. The experimental corpus consisted of a KPI dataset containing data from 70 cells of a live mobile network. For each cell, 12 KPIs were collected every hour for four months, from 11/15/2011 to 03/19/2012. The KPIs have different characteristics; some of them, such as downlink or uplink data volume or throughput, are measurements of user traffic utilization; while others, such as drop-call rate and successful call-setup rate, are measurements of call control parameters.

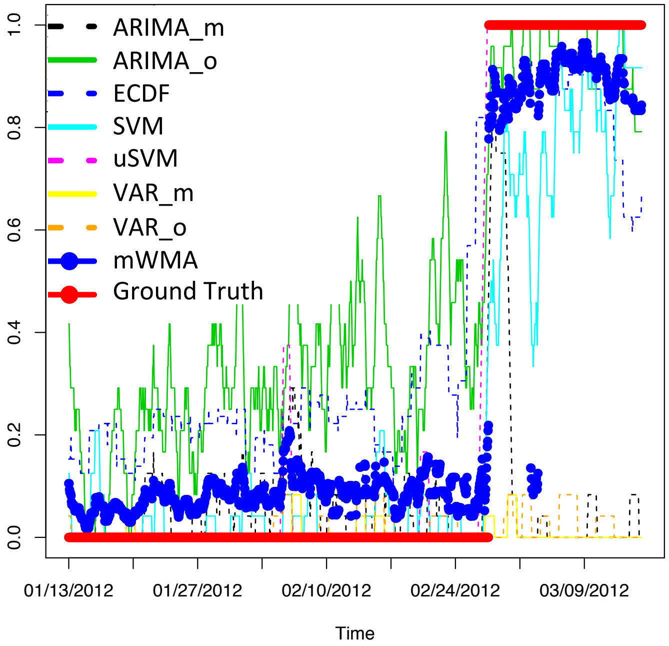

Fig. 2. The output KPI degradation levels generated by the ensemble method for a given cell and call control KPI are marked with blue circles, while red represents the manually generated labels. The remaining series represent the KPI degradation levels generated by the univariate and multivariate methods ($\tau = 0.5$ and $\beta = 0.8$)

The experimental dataset had no associated ground truth. To address this limitation, labels were manually generated to indicate whether the data represented a real problem or not, based on engineering knowledge applied to KPI-data visual inspection.

The pool of models was trained on the first 912 hours of data, and the ensemble method was trained on the next 500 hours). The remainder of the dataset was used to make the ensemble prediction based on the learned weights. The parameters were set to $\tau = 0.5$ and $\beta = 0.8$.

Figure 2 presents the KPI degradation levels generated by the ensemble methods (modified WMA depicted as mWMA) as well as the univariate and multivariate methods.

The two metrics used for the performance evaluation were:

- False Positive Rate (FPR) defined as the percentage of normal data deemed as abnormal by the detector
- Detection Rate (DR) defined as the percentage of abnormal data deemed as abnormal by the detector.

Figure 3 presents the Receiver Operating Characteristic (ROC) [12] curve for all the methods (the detection and false

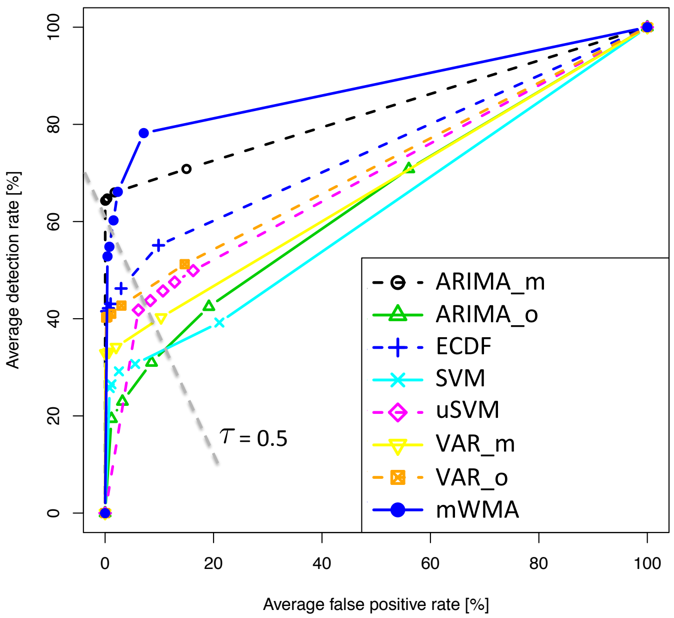

Fig. 3. ROC curves for all individual methods and the ensemble method

positive rates were computed as an average across all the 70 cells analyzed) and it illustrates well that the ensemble method (mWMA) exhibits the best performance, confirming our hypothesis.

IV. Related Work

The proposed framework aims to detect partial and complete degradations in cell-service performance. Previous research addressed the cell-outage detection [2] and cell-outage compensation [3] concepts. For the problem of cell-outage detection, Mueller et al. [2] proposed a detection mechanism that uses Neighbor Cell List (NCL) reports. Compared to our work, Muller's approach was limited to only catatonic-cell detection, while not every isolated node reflected an outage situation.

Another approach for estimating failures in cellular networks was proposed by Coluccia et al. [13] to analyze events at different levels: transmission of IP packets, transport and application layer communication establishment, user level session activation, and control-plane procedures.

D'Alconzo et al. [14] proposed an anomaly detection algorithm for 3G cellular networks that detects events that might put the stability and performance of the network at risk.

More recently, detection of general anomalies has also been addressed [4], [5], [6]. However, to the best of our knowledge, our approach is the first to employ an adaptive ensemble method that copes with concept drift.

V. Conclusions and Future Work

This paper proposed a novel ensemble method for modeling cell behavior that builds adaptive models and uses the intrinsic characteristics of the environment where the models are created to improve its performance. The design was implemented and applied to a dataset consisting of KPI data collected from a real operational cell network. The experimental results indicate that our system provides significant detection performance improvements over stand-alone univariate and multivariate methods.

We are currently planning experimental evaluation of our cell anomaly detection method in a network operator setting. Additional work is needed to integrate our detection component with a diagnosis engine that combines the detector output with other information sources to assist operators in determining the cause of a detected anomaly. These results also serve as the foundation for research in other areas of network operation, specifically to evaluate the impact of configuration changes on critical measures of network performance.

Acknowledgment

We thank Lauri Oksanen, Kari Aaltonen, Richard Fehlmann, Christoph Frenzel, Péter Szilágyi, Michael Freed, Ken Nitz, 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] C. M. Mueller, M. Kaschub, C. Blankenhorn, and S. Wanke, "A Cell Outage Detection Algorithm Using Neighbor Cell List Reports," International Workshop on Self-Organizing Systems, 2008.

[3] M. Amirijoo, L. Jorguseski, R. Litjens, and L.C. Schmelz, "Cell Outage Compensation in LTE Networks: Algorithms and Performance Assessment," 2011 IEEE 73rd Vehicular Technology Conference (VTC Spring), 15–18 May 2011.

[4] A. Bouillard, A. Junier and B. Ronot, "Hidden Anomaly Detection in Telecommunication Networks," International Conference on Network and Service Management (CNSM), Las Vegas, NV, October 2012.

[5] P. Szilágyi and S. Nováczki, "An Automatic Detection and Diagnosis Framework For Mobile Communication Systems," IEEE Transactions on Network and Service Management, 2012.

[6] S. Nováczki, "An Improved Anomaly Detection and Diagnosis Framework for Mobile Network Operators," 9th International Conference on Design of Reliable Communication Networks (DRCN 2013), Budapest, March 2013.

[7] S. Rüping, "SVM Kernels for Time Series Analysis," In R. Klinkenberg *et al.* (eds.), LLWA 01 - Tagungsband der GI-Workshop-Woche Lernen - Lehren - Wissen - Adaptivität, Forschungsberichte des Fachbereichs Informatik der Universität Dortmund, pp. 43-50, Dortmund, Germany, 2001.

[8] J. Ma and S. Perkins, "Time-Series Novelty Detection Using One-Class Support Vector Machines," Neural Networks, 2003.

[9] R. B. Cleveland, W. S. Cleveland, J. E. McRae, and I. Terpenning, "STL: A Seasonal-Trend Decomposition Procedure Based on Loess," Journal of Official Statistics, Vol. 6, No. 1, 1990.

[10] B. Pfaff, "VAR, SVAR and SVEC Models: Implementation Within R Package vars," Journal of Statistical Software, Vol. 27, Issue 4, 2008.

[11] N. Littlestone and M.K. Warmuth, "The Weighted Majority Algorithm," Inf. Comput. 108, 2, 1994.

[12] D. M. Green and J. A. Swets, Signal Detection Theory and Psychophysics. New York, NY: John Wiley and Sons Inc. ISBN 0-471-32420-5, 1966.

[13] A. Coluccia, F. Ricciato, and P. Romirer-Maierhofer, "Bayesian Estimation of Network-Wide Mean Failure Probability in 3G Cellular Networks," In PERFORM, Vol. 6821, Springer (2010), pp. 167-178.

[14] A. D'Alconzo, A. Coluccia, F. Ricciato, and P. Romirer-Maierhofer, "A Distribution-Based Approach to Anomaly Detection and Application to 3G Mobile Traffic," Global Telecommunications Conference (GLOBECOM) 2009.