

# Social-Aware Forecasting for Cellular Networks Metrics

Javier Villegas<sup>1</sup>, Eduardo Baena<sup>1</sup>, Sergio Fortes<sup>1</sup>, *Member, IEEE*, and Raquel Barco<sup>1</sup>

**Abstract**—It has been long established that crowds generated by social events (e.g., sports matches, parades, fairs...) produce a high impact on cellular network service. However, to estimate such an impact, it is necessary to use data sources classically outside the mobile operator control. In this way, and following a social-aware approach, the forecasting mechanisms should be able to combine both social and network information to obtain reliable predictions. To this end, the present work develops a complete system for its use in the prediction of cellular metrics (e.g., connections, throughput...). The performance of the proposed solution is evaluated in a real cellular network, showing the capabilities of the approach to provide accurate forecasting.

**Index Terms**—Cellular communications, self-organizing networks, social events, network intelligence, forecasting.

## I. INTRODUCTION

NEW and increasingly demanding applications are taking the management of 5G networks to new heights of complexity. Particularly when there are large concentrations of users (sports matches, concerts, parades, etc.), whether planned as social events or spontaneously set, it becomes challenging for the cellular network operators to maintain the quality of the service and attend to all possible issues. Unexpected crowds of users commonly cause network overloads and, thus, degradations in the network service. In this context, the automation of cellular network management proposed under the paradigm of the Self Organizing Networks (SON) becomes growingly relevant. Hence, the availability of SON mechanisms for self-healing and self-optimization makes them a critical element for current and next-generation cellular networks to maximize the utilization of the available resources and meet demand requirements.

Classic approaches to manage performance degradations in cellular networks have been typically based purely on cellular metrics such as alarms, counters, KPIs (Key-Performance Indicators), user traces, etc. It has been only recently that other sources of information, such as application-specific indicators [1] and context variables (e.g. the users' location [2]) have been considered as key input drivers for the operation, administration and management (OAM) tasks of the cellular

network. Regarding location as context information, in [3], Li *et al.* proposed a fog-based latency forecasting system for Internet of Things networks, helping to minimize the delay in those scenarios. Moreover, in [4], Li *et al.* presented a virtualization enabled fog computing framework, where data is computed in near nodes, and, reducing load and costs in cellular networks.

On the other hand, although the impact of crowds in the network is very high, the identification of the social events causing them, if present, is typically not automated and reactive in nature [5]. Forecasting mechanisms able to predict the impact of these events into the cellular metrics are deemed indispensable to support proactive resource allocation and optimization mechanisms able to avoid or compensate their expected impact in the network.

This makes necessary to incorporate the social dimension (inherent to the service demand) into the analysis of network performance. In this sense, in recent years, a few works have proposed the use of this social context in the field of prediction with various management applications, leading to what can be defined as “social-aware” systems.

In this area of social-awareness, most works have focused on reactive mechanisms. In this way, in [6], Trinh *et al.* proposed a framework to detect anomalies using an LSTM (Long Short-Term Memory) neural network model that associate abnormal values in the network metrics to social events. Furthermore, the authors in [7] proposed another anomaly detection system based on the prediction error when traffic is higher compared to its normal behavior. Similar techniques have been proposed where information from Social networks (i.e., Twitter) is combined with historical cellular data to detect coverage holes, estimate user demand, or to plan new base stations [8], [9]. However, the proposed techniques are typically associated with a specific metric and lack the ability to predict future metric values. Also, Twitter-based geolocated information is scarce and alone will typically not be enough for social events characterization, making necessary the use of other information sources.

Beyond specific mechanisms, social-awareness has been also the focus of wider management frameworks. In [5], Fortes *et al.* developed a system to associate past anomalies to previous social events, allowing to identify its root cause. Nevertheless, the proposed approach does not provide techniques for the forecasting of future anomalies. Furthermore, in [10], Pintér *et al.* proposed a framework that allows to understand mobility patterns during large social events using logs obtained from the LTE (Long Term Evolution) network. Nevertheless, such model relied exclusively on previously captured measurements and highlighted the difficulty of using online social data. Accordingly, the work in [11] surveyed the role that context information plays towards network optimization through the knowledge of the future and predictions, providing a summary

Manuscript received February 16, 2021; accepted March 4, 2021. Date of publication March 12, 2021; date of current version June 10, 2021. This work has been partially funded by Junta de Andalucía and ERDF in the framework of the projects IDADE-5G (UMA18-FEDERJA-201, “Programa Operativo FEDER Andalucía 2014-2020”) and SMART - “Sistema de Monitorización y análisis Automático de Redes de Telecomunicación” (“Incentivos a los agentes públicos del Sistema Andaluz del Conocimiento, para la realización de proyectos de I+D+i., en el ámbito del Plan Andaluz de Investigación, Desarrollo e Innovación PAIDI 2020”). It has been also performed in the framework of the Horizon 2020 project LOCUS (ICT-871249) receiving funds from the European Union. The associate editor coordinating the review of this letter and approving it for publication was S. Yu. (*Corresponding author: Sergio Fortes.*)

The authors are with the Departamento de Ingeniería de Comunicaciones, Andalucía Tech, Universidad de Málaga at Teatinos, 29071 Málaga, Spain (e-mail: jvc@ic.uma.es; ebm@ic.uma.es; sfr@ic.uma.es; rbm@ic.uma.es).

Digital Object Identifier 10.1109/LCOMM.2021.3065812

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.

For more information, see <https://creativecommons.org/licenses/by-nc-nd/4.0/>

of the most promising techniques in the field. Nevertheless, long-term predictions that could be applicable regardless of the chosen metric are not identified, rather each cellular metric is individually studied.

Finally, a recent survey [12] identified the need for automatic data-driven and proactive network optimization systems, i.e., with the network capacity to predict and act in advance. Here the use of social data has been linked with caching decisions, only suggesting as future work the aggregation of social data to the cellular metrics and the application of multi-layer neural networks.

Considering the above, a shortage of developments on cellular metrics forecasting making use of social data is identified. Going beyond previous approaches, the main contributions of this letter include the definition of a novel framework for the prediction of cellular network metrics in a generalized manner, using both social and cellular data. For this, the use of social events' start dates alongside cellular data is proposed. To allow this, a novel approach of stacked *nonlinear autoregressive exogenous models* (NARX) is adopted, being specifically defined in order to allow the use of the social event information to improve the metrics forecasting. Moreover, the methodology for its hyperparameters estimation, training, and forecasting application is established, testing it based on real cellular data.

In this way, this letter is organized as follows. Section II presents the main algorithms used for forecasting. In Section III, the novel system constructed in order to support these mechanisms is presented, including a detailed description of its functions and variables. Then, in Section IV, the predictions provided by the system are evaluated using data from a real cellular network and compared to a non-social based approach. Finally, Section V summarizes the main conclusions.

## II. NARX NEURAL PREDICTION

To predict future metrics, a NARX model is proposed. NARX models use the past samples of the metric,  $y[n]$ , and the current and future values of an exogenous input,  $x[n]$ , to calculate future estimated values  $y'[n]$  of  $y[n]$  [13].

To implement this model, a multilayer perceptron (MLP), a type of feedforward neural network approach, is proposed [14]. This is expressed in Equation (1):

$$y'[n]=\Psi(y[n-1]\dots y[n-n_y],x[n-1]\dots x[n-n_x],b[n]), \quad (1)$$

where  $\Psi$  is the function of the model where  $n_y$  and  $n_x$  are the number of past samples of the variables  $x[n]$  and  $y[n]$ , respectively, taken as inputs.  $b[n]$  represents *bias values*, a constant added to shift the output of a neuron without varying the weights of the inputs.

Following a similar approach to the one described (for general applications) in [15], in the proposed system, a number  $k$  of these predictor models are stacked in order to increase the robustness and reliability of the predictions by combining their output using the median value of the predictions. Therefore, if a predictor fails, the resulting forecast is still correct, thanks to other predictors.

The structure of this model is shown Figure 1 where two configurations are defined: the first one is in *open-loop* and

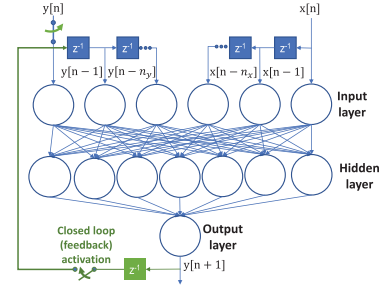


Fig. 1. Proposed MLP structure.

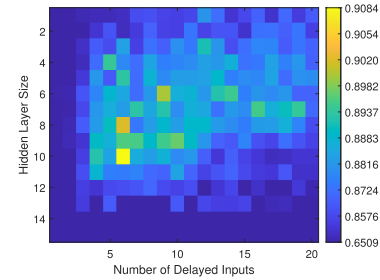


Fig. 2. Hyperparameters median correlation.

it is the one used for training. Once trained, the *closed-loop* architecture is used for the prediction.

In order to define the optimal values for the *hidden layer size* and the *number of delayed inputs*, a hyperparameters selection module is defined. This estimates these parameters of the network by testing all possible combinations of these two parameters within a certain value range. This process is presented in Figure 2 for the *dl\_user\_throughput* metric to be used as input and further described in the evaluation section. The median of the correlation between the predictions and the real values of the metric is calculated for 100 experiments for each of the possible combinations of hidden layer size and number of delayed inputs. From this, the pair of hyperparameters with the highest median correlation is selected, in this case a hidden layer size of 10 and 6 delayed inputs.

The performance of the provided forecast will be also related to the dependency of each of the metrics in respect to the number of users and their demand. Hence, SAFe provides an indication of the expected quality of the forecasts for each metric based on the correlation results obtained for the training set during the hyperparameters selection process.

## III. PROPOSED SYSTEM

With the proposed NARX forecasting model and social data as a key elements, a Social-Aware Forecasting framework for cellular network metrics (SAFe) is proposed. SAFe goes beyond previous social acquisition system, such as the one proposed in [5] by adding a complete set of functionalities aiming to provide the necessary inputs for the forecasting mechanisms. As shown in Figure 3, SAFe functionalities are going to be described in detail in the next subsections.

### A. Social-Aware OAM Support Block

This module is based on the system defined in [5] and it is charge of providing the social data required by the framework. As shown in Figure 3, it includes a set of multiple steps.

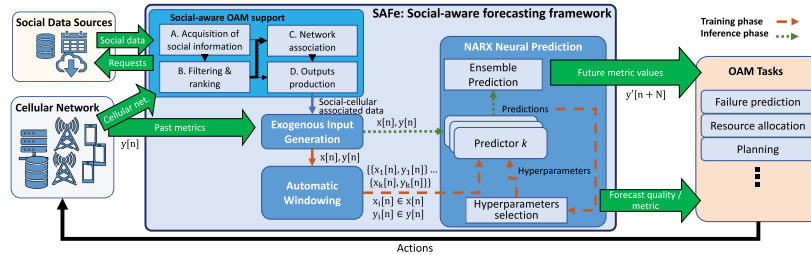


Fig. 3. High-level diagram of the social-aware forecasting system.

Firstly, the *acquisition of social information* block gathers social data (mainly event start date, venue location coordinates, type) from calendar sources (e.g. databases, events aggregators). This information is *filtered and ranked* (block B) by the relevance of the events eliminating those where the available information is not complete or go outside of the geographical area under analysis [5].

From this, block C, *network association* estimates the set of venues (e.g. stadiums, concert halls, parks...) and events that could likely have impacted each cell. This is estimated by selecting for each cell those events that are closer and in the proper bearing to the center of its radiation pattern. This is directly calculated based on the known geographical position of the sites and the azimuth of their sectors. Additional filtering can be performed by correlating past cell metric values with events from a specific venue or area [5], discarding events from low-impacting venues and keeping those of high impact (e.g. restaurant events would not generate a high impact in the network while those coming from a stadium will typically do). This information is then provided as an output by block D in order to feed the posterior modules of the SAFE system.

### B. Exogenous Input Generation

Once the events associated with degradations have been identified, and the venues where those took place have been selected, the exogenous input,  $x[n]$ , for the NARX model (as presented in Equation (1)) is generated. This is created to encompass in a time-series manner the information about the start of each event that can impact the cell under analysis, in order to be able to associate events to the metric being predicted. To do so,  $x[n]$  generation is defined by a binary signal encompassing the time instant when an event starts (valued as '1') or not ('0'). Since the events can be known in advance, this exogenous input will contain past values as well as future values.

### C. Automatic Windowing

In order to apply the NARX model, raw cellular data,  $y[n]$ , is typically not applicable as anomalies associated with social events might overlap with each other, for example, when two successive events occurs within few hours. To overcome this a *windowing* algorithm is defined in order to separate metrics into different segments associated with an increase and a posterior decrease on demand-related values as they could be associated with an event, as it can be observed in Figure 4. This process consists of dividing the data into segments that keep a pattern formed by an initial increase followed by a decrease in its values associated with user demand. These

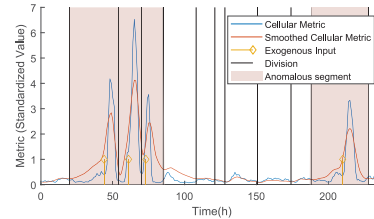


Fig. 4. Example of raw metric data.

segments can be reordered, obtaining a new time series and, therefore, allowing to experiment with different scenarios using the same dataset. To delimit these event segments, the proposed segmentation is based on the zero crosses (ZCs) of the derivative function of the metric. However, there are periods in which too many ZCs would appear, even if no significant change in the metric  $y[n]$  has taken place. To avoid this issue, a smoothed “envelope” of the metric is obtained from the Hilbert transform. Based on the ZCs of the derivative of this transform, the segments are calculated. Figure 4 shows this process where the filtered Hilbert envelope is shown for an example metric (the number of active connections).

In order to classify if each of these segments are *anomalous* and probably affected by an event, the range of metric values considered as normal (without the presence of an event) is established by low ( $th_{low}$ ) and high ( $th_{high}$ ) thresholds. Such values, in turn, are calculated from the mean ( $\mu$ ) and variance ( $\sigma$ ) as follows:  $th_{low} = \mu(y) - Tol * \sigma(y)$  and  $th_{high} = \mu(y) + Tol * \sigma(y)$ . In this context, the tolerance ( $Tol$ ) establishes the severity level of the threshold. This is calculated automatically by the expression:  $Tol = \sqrt{2 \ln(|y|)}$ , being  $|y|$  the number of samples of the metric in a segment.

Lastly, the dataset is divided into as many subsets as anomaly segments. The event anomalies are separated by introducing non-anomalous segments between them. Figure 5 shows an example of the resulting *separated* metric from the portions obtained from the original data. The different reorders of these segments allow for the posterior cross-validation of the forecasting mechanisms.

## IV. EVALUATION

The evaluation of SAFE has been carried out using data from a real LTE network. Here, the focus is on two key example cells in the same area described in [5]. Firstly, *Cell\_A* is a cell directly covering a big stadium affected by several social events. Secondly, *Cell\_B* is located near the stadium entrance surroundings. The dataset contains 238 hourly samples from *Cell\_A* and 571 from *Cell\_B*. For these, prediction accuracy has been tested up to 24 steps (hours) ahead, using a system



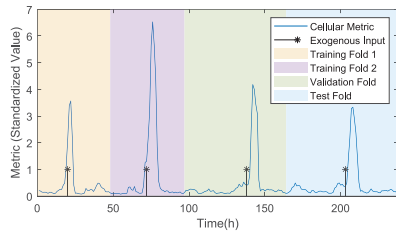


Fig. 5. Example of metric data divided in folds.

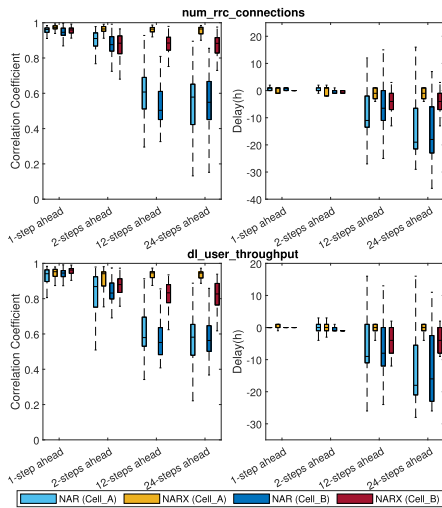


Fig. 6. Correlation coefficients and delays obtained.

with 3 redundant predictors and performing 100 experiments for each metric.

Since forecasts may be time-shifted with respect to the real values, Mean Absolute Error (MSE) or Root Mean Squared Error (RMSE) are measures that would give very high error levels even for small-time differences with similar values. It has to be considered that small shifts in the start of the events will not highly impact management tasks dedicated to compensate their effects, being more relevant that the predicted maximum metric values and temporal profile are close to the real ones.

Considering this, the cross-correlation between the estimated and the predicted values is applied as the key figure of merit of the proposed framework. As a complement to this, the maximum values of the correlation and the delay between both are calculated. Also, to improve the display of the results, correlation is presented as a normalized coefficient. In Figure 6 the results for *Cell\_A* and *Cell\_B* are plotted for two key metrics: *num\_rrc\_connections* (i.e. number of user connections) and *dl\_user\_throughput* (i.e. user download speed). In order to prove the advantage of the proposed method against traditional approaches, a *nonlinear autoregressive model* (NAR), which corresponds to the proposed model without exogenous input, is used as a baseline to compare the NARX results.

It can be observed that the median value of NARX correlation between real metrics and predictions for all the calculated “steps ahead” values and both cells is above 0.8, providing an accurate prediction, even in the long-term. Nevertheless, the graph shows a higher correlation degree for the cell directly

covering the stadium. Regarding the delay, there is almost no difference between metrics, achieving the proposed NARX system low median delay values (below 5 hours).

## V. CONCLUSION

This work has proposed a completely novel framework to improve the prediction of any relevant network metric that incorporates the relationship between social events and crowd gatherings and their impact on the cellular network. Results based on real cellular network data have proved the effectiveness of the method. From the analysis of the predictions made for the different metrics, it is possible to conclude the high accuracy of the system to generate forecasts using only social information and data from past events, even in the long-term.

Future works will use such predictions as inputs for SON mechanisms in order to improve network performance. Moreover, the application of the proposed scheme in the context of SARS-CoV-2 and general pandemic scenarios (where crowds and mobility are restricted) will be further analyzed.

## REFERENCES

- [1] C. Baena, S. Fortes, E. Baena, and R. Barco, “Estimation of video streaming KQIs for radio access negotiation in network slicing scenarios,” *IEEE Commun. Lett.*, vol. 24, no. 6, pp. 1304–1307, Jun. 2020.
- [2] S. Fortes, A. Aguilar-García, R. Barco, F. Barba, J. Fernández-luque, and A. Fernández-Durán, “Management architecture for location-aware self-organizing LTE/LTE—A small cell networks,” *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 294–302, Jan. 2015.
- [3] J. Li, T. Zhang, J. Jin, Y. Yang, D. Yuan, and L. Gao, “Latency estimation for fog-based Internet of Things,” in *Proc. 27th Int. Telecommun. Netw. Appl. Conf. (ITNAC)*, Nov. 2017, pp. 1–6.
- [4] J. Li, J. Jin, D. Yuan, and H. Zhang, “Virtual fog: A virtualization enabled fog computing framework for Internet of Things,” *IEEE Internet Things J.*, vol. 5, no. 1, pp. 121–131, Feb. 2018.
- [5] S. Fortes, D. Palacios, I. Serrano, and R. Barco, “Applying social event data for the management of cellular networks,” *IEEE Commun. Mag.*, vol. 56, no. 11, pp. 36–43, Nov. 2018.
- [6] H. D. Trinh, L. Giupponi, and P. Dini, “Urban anomaly detection by processing mobile traffic traces with LSTM neural networks,” in *Proc. 16th Annu. IEEE Int. Conf. Sens., Commun., Netw. (SECON)*, Jun. 2019, pp. 1–8.
- [7] H. D. Trinh, E. Zeydan, L. Giupponi, and P. Dini, “Detecting mobile traffic anomalies through physical control channel fingerprinting: A deep semi-supervised approach,” *IEEE Access*, vol. 7, pp. 152187–152201, 2019.
- [8] H. Klessig, H. Kuntzschmann, L. Scheuven, B. Almeroth, P. Schulz, and G. Fettweis, “Twitter as a source for spatial traffic information in big data-enabled self-organizing networks,” in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2017, pp. 1–5.
- [9] J. L. Bejarano-Luque, M. Fernandez-Navarro, M. Toril, A. J. Garcia, and S. Luna-Ramirez, “A context-aware data-driven algorithm for small cell site selection in cellular networks,” *IEEE Access*, vol. 8, pp. 105335–105350, 2020.
- [10] G. Pinter, L. Nadai, and I. Felde, “Analysis of mobility patterns during a large social event,” in *Proc. IEEE 16th Int. Symp. Intell. Syst. Informat. (SISY)*, Sep. 2018, pp. 000339–000344.
- [11] N. Bui, M. Cesana, S. A. Hosseini, Q. Liao, I. Malanchini, and J. Widmer, “A survey of anticipatory mobile networking: Context-based classification, prediction methodologies, and optimization techniques,” *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1790–1821, 3rd Quart., 2017.
- [12] B. Ma, W. Guo, and J. Zhang, “A survey of online data-driven proactive 5G network optimisation using machine learning,” *IEEE Access*, vol. 8, pp. 35606–35637, 2020.
- [13] J. M. P. Menezes and G. A. Barreto, “Long-term time series prediction with the NARX network: An empirical evaluation,” *Neurocomputing*, vol. 71, nos. 16–18, pp. 3335–3343, Oct. 2008.
- [14] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, Eds., “Deep learning,” in *Data Mining*, 4th ed. San Mateo, CA, USA: Morgan Kaufmann, 2017, ch. 10, pp. 417–466.
- [15] P. Bühlmann, *Bagging, Boosting and Ensemble Methods*. Berlin, Germany: Springer, 2012, pp. 985–1022.