International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 28 Issue 6 – JUNE 2021.

PREDICTION OF TRAFFIC FORECASTING USING ANDROID APPLICATION

Mrs.S.Sivaraman[#1] and Ms.C.Ambika[*2] and Ms.B.Bavya[*3] and Ms.G.Bhavani[*4]

Assistant Professor, Department of Computer Science and Engineering, Vivekanandha College of Engineering forWomen, Tiruchengode, ,Tamilnadu.,India

* Department of Computer Science and Engineering, Vivekanandha College of Engineering for Women, Tiruchengode, ,Tamilnadu.,India

ABSTRACT- We consider the issue of determining high recurrence tested versatile cell traffic beginning from a lower recurrence examined time arrangement. We utilize a dataset of genuine downlink/uplink traffic follows acquired from a versatile cell arrange and apply various techniques for performing estimates at various testing frequencies. Through broad assessment we show that such kind of guaging is conceivable and now and again is additionally ready to beat figure results got beginning straightforwardly from the high recurrence time arrangement. The results of this work can be utilized for a few situations of intellectual systems administration, including expectation of information traffic demands in explicit areas, just as for information stockpiling streamlining and improvement of BBU bunching errands.

1. INTRODUCTION

The forthcoming fifth era (5G) of portable organizations will uphold a phenomenal development in the quantity of associated gadgets and the volume of information traffic, because of a progression of advancement in the manner the whole organization engineering is planned and worked. Specifically, future organizations are relied upon to firmly depend on network work virtualization (NFV) and programming characterized organizing (SDN) at the center, while the radio access part will exploit creative arrangements, for example, Multi-access Edge figuring (MEC) and Cloud-RAN (C-RAN). The last depends on the accumulation of all Base Stations computational assets in an incorporated baseband unit (BBU) pool, which is associated with a thickly dispersed arrangement of Remote Radio Heads (RRH) through high limit joins. Contrasted with conventional models, the CRAN approach can ideally allot assets in a unique manner to the individual RRH, whose equipment configuration is a lot less difficult than heritage LTE eNodeBs. This methodology permits to expand network asset usage, decline obstruction and simultaneously bringing down energy utilization and by and large equipment costs [1]. In this situation, it is imagined that enormous information examination and AI/information driven procedures will assume a significant part in all periods of the cycle. Specifically, guaging future downlink/uplink traffic requests is critical for C-RAN activities: for instance, when the BBU capacities are virtualized utilizing virtual holders (e.g., Docker) which can be increased and down powerfully, the instatement time needed for such level scaling1 1https://kubernetes.io/docs/errands/run-application/even unit autoscale tasks is in the request for 1-5 minutes on current equipment [2]. In this manner, it is essential to gauge transient traffic requests to have the option to opportune oblige the required BBU assets. Regardless, to restrict the measure of capacity assets required, network administrators total their datasets either transiently (e.g., figuring hourly or day by day totals of traffic requests for each eNodeB) or spatially (e.g., collecting information from a few eNodeBs in a solitary rush hour gridlock arrangement). Notwithstanding, while such accumulations represent a cutoff on the required stockpiling assets, they have the disadvantage of potentially impeding fine-grained estimating required for future C-RAN tasks. The goal of this paper is to evaluate whether is feasible to reuse transiently accumulated traffic information to perform momentary anticipating. Specifically, we center around uplink/downlink video traffic request follows got from a genuine country-wide versatile cell arrange and dissect the exhibition of present moment/high-recurrence guaging (e.g., 5 minutes-ahead forecasts) beginning from input information which is examined at a lower recurrence (e.g., hourly). We use AI techniques to play out the gauge tasks and remark on the presentation acquired contrasted with different baselines strategies ordinarily utilized in the writing. The results of this work can be utilized for a few situations of psychological systems administration, including expectation of information traffic demands in explicit areas (eNodeBs), just as for information stockpiling streamlining and improvement of BBU bunching undertakings.

Downlink traffic forecast

NRMSE obtained for the case of downlink traffic. Let's focus first on the baseline methods, that is forecasting methods applied to the 5-minutes sampled time series.

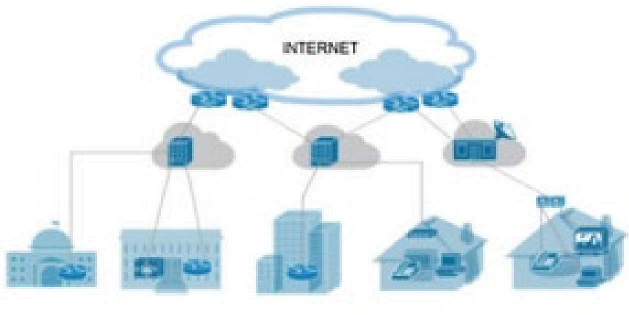

Fig 1 : eNodeBs Internet

The last value approach is the simplest method available and allows to obtain a NRMSE of approximately 0.9. The ARIMA baseline allows to increase the performance at the cost of a slightly higher number of input samples in the model: a grid search for the optimal AR parameters return p = 7 as best value, with a corresponding NRMSE of less than 0.8. Baselines using machine learning obtain the best results, with Neural Networks whose performances increase as the number of input samples increases, while Linear Regression models behave in the opposite way (e.g., the performance decrease as the number of input samples increase). This is explained observing that while the Neural Network used in this work has a fixed amount of parameters to be estimated (i.e. one single hidden layer with 32 nodes), the Linear Regression model has a number of parameters that grows with the number of input samples. Therefore, given the limited amount of training data (2 weeks), when increasing N the number of observations available for training Linear Regression models decreases, and this explains the performance degradation. Moving to the results obtained by Time Series Distribution (TSD) and Direct Forecasting (DF) approaches, that is forecasting a high-frequency series starting from a low-frequency one, we can observe how the performance of such methods surprisingly approach the baselines, except when used in conjunction with ARIMA, for which baseline model behaves better. In both approaches, again the Neural Network methodologies increase their performance as the number of input samples increases. For what concerns linear regression models, again we notice how the performance tend to decrease as the number of input samples increases (for the same reason explained before). Interestingly, the approach based on Time Series Distribution allows to limit the performance degradation compared to the one using Direct Forecasting when using linear regression models. Similar comments can be made for the case of NUERMSE, where the surprisingly good performance of methods such as TSD-MLR or DF-NN is highlighted. It can be seen that, especially for high number of input samples, such methods obtain performances which are even better than baseline methods. That means that it is possible to forecast high-frequency traffic starting from lowfrequency time series better than what is possible to do starting from the corresponding high-frequency series.

Uplink traffic forecast

We also study the case of uplink traffic. This is particularly interesting for C-RAN operations, as many works proposed a centralized processing of HARQ and FEC schemes, which are characterized by tight time requirements. Being able to forecast accurately uplink traffic is therefore crucial to understand how many resources will be needed in the cloud-RAN for such type of processing. NRMSE performance obtained for the different proposed methods. In general we can observe that the performances are worse compared to the downlink case: this is due to the high variability of uplink data traces, whose behavior is much more unpredictable than downlink traffic. We can observe that the last value baseline method isn't able to perform accurate forecast (NRMSE greater than 1), while ARIMA and machinelearning based baselines perform just better than the average prediction method. Similarly to the downlink case, we can observe that: i) Neural Networks improve their performances as the number of input samples increases, while this is not true for Linear Regression and ii) surprisingly, TSD and DF methods in combination with Neural Networks outperform the baseline approaches, while TSD-ARIMA is outperformed by ARIMA baseline model. This means that also in the uplink case using a low-frequency traffic time series can be used to forecast higher-frequency traffic samples, even though with performance which are in general worse compared to the downlink scenario.

Theoretical Background

The use of ANN is encapsulated in a wide area of applications on the grounds of its generalization capability to unforeseen situations. ANN is trained on a set of inputs to reach a specific target output using an appropriate learning algorithm until the outcomes of the network tend to match the given targets. This way the parameters of the network are selected, which are able to deduce the nonlinear relation between given inputs and outputs. In order to increase the efficiency of ANN, its computational capacity requires more neurons per layer or increasing the number of hidden layers. This results in the deep architecture of neural computational model, in which the intermediate layers or hidden layers perform as multiple layers of abstraction. Massive research is taking place as far as deep architectures of ANN are concerned. Deep ANNs are achieving encouraging improvements over previous shallow ANN architectures [6]. Deep ANNs contain numerous levels of non-linearities depending on the depth of hidden layers. The deep hierarchical architecture allows them to efficiently represent highly nonlinear patterns and highly-varying functional abstractions. Although, it was not clear how to train such deep networks, as the random initialization of network parameters appears to often get stuck in poor solutions [7]. In [8] a learning algorithm was introduced that greedily trains, one a layer of network at a time. Deep architectures can be formed either by stacking autoencoders or Deep Belief Networks (DBNs) of RBMs [8]. DBN is the most common and effective approach among all deep learning models. DBN is formed by stacking of RBMs. RBM is a bipartite graphical, energy based, probabilistic model and it is interpreted as stochastic neural network. RBM relies on two layer structure comprising on visible and hidden nodes as shown in Fig. 14.2. These nodes are binary units which means h and v ϵ {0, 1}. These nodes are conditionally independent given one another. Therefore, the probabilistic representation of these hidden and visible nodes with sigmoid activation function i.e. σð Þx = 1/1 + e−x ð Þ is given as in Eqs. (1) and (2). Wij is a weight associated with the edge between units Vj and hi. Whereas bj and ci are real valued bias terms associated with the jth visible and the ith hidden variable, respectively. Sampling in an RBM is obtained by running a Markov chain to convergence, using Gibbs sampling.

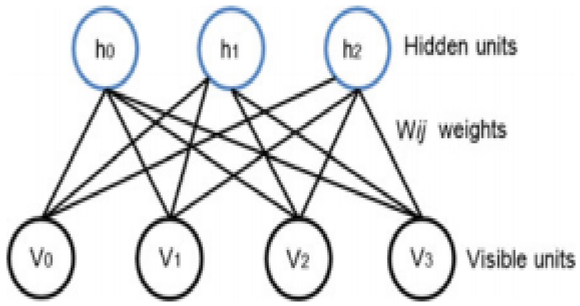

Fig 2 : Vehicle Visible Units

2.PROBLEM STATEMENT :

Since a record corresponds to an L2 connection, a record finishes when an L2 device at the communication path detects a certain period of silence and disconnects the communication. On the other hand, immediately after an application on L2 ends, another application begins to run on the same L2 connection, these applications are logged into the same record. (a) shows that two application sessions are combined into a record due to a short SIT. The value of SIT determines whether or not two consecutive sessions are combined and the values depends on the expiration time at L2. Fig.1 (b) exhibits SSMR, where a session is split into three records. This phenomenon occurs when handover take place to change the used cell. As stated above, the relationship between sessions and records are not necessarily a one-toone mapping. Although we consider complicated non-one-toone mapping later. At the moment, we first examine only one-to-one mapping to see the possibility of estimating used applications by only using CDRs. It should be noted that Android phones generate background traffic without the user's control. The background traffic occurring in SIT can result in an MSCR. In addition, a session solely with background traffic worsens the accuracy of estimating the application type. Therefore, we need to identify how often and how much background traffic is generated. Our experiment using a Wi-Fi access point shows that LINE generates data with less 250 bytes every 30 seconds and 200 seconds when it runs and when it does not run, respectively.

4. LITERATURE REVIEW

Cloud radio access network (c-ran): a primer

Cloud Radio Access Network (C-RAN) is expected to be a candidate of next generation access network techniques that can solve operators' puzzle. In this article, on the basis of a general survey of C-RAN, we

5.SYSTEM ARCHITECTURE

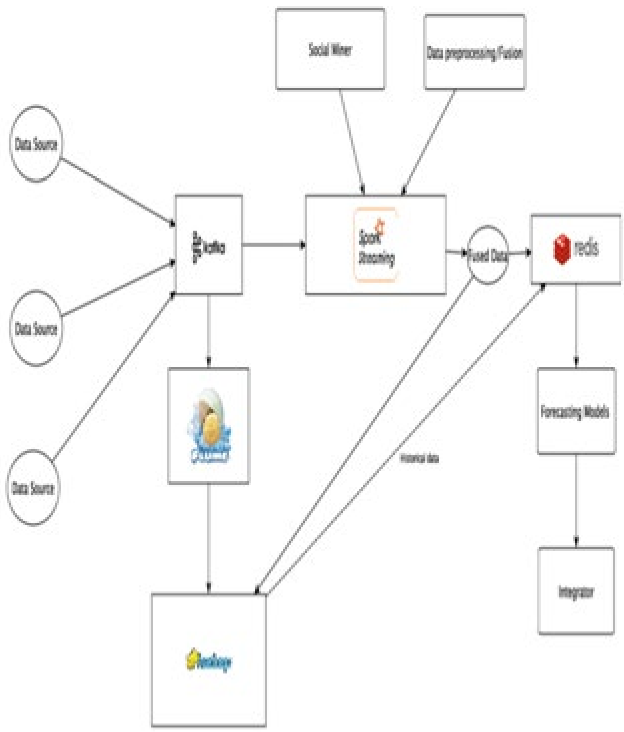

Fig : 3 System Architecture

present a novel logical structure of C-RAN that consistsof a physical plane, a control plane, and a service plane. Compared to traditional architecture, the proposed C-RAN architecture emphasizes the notion of service cloud, service-oriented resource scheduling and management, thus it facilitates the utilization of new communication and computer techniques. With the extensive computation resource offered by the cloud platform, a coordinated user scheduling algorithm and parallel optimum precoding scheme are proposed, which can achieve better performance. The proposed scheme opens another door to design new algorithms matching well with C-RAN architecture, instead of only migrating existing algorithms from traditional architecture to C-RAN.

6.EXISTING SYSTEM

The existing works are described by considering five criteria; the used video preprocessing tools, counting approach, whether the algorithm is real-time or not (R-T), validation data size, and system accuracy. These factors motivate us to go further by proposing a new application that is real time, simple to use, automatic, and validated on long-lasting videos from different real scenarios, including, different smartphones, video duration, geographical position, intrinsic and extrinsic camera characteristics, and weather conditions.

7. PROPOSED METHODOLOGY

The developed app can be used by many public agencies, private sectors, governments for traffic flow analysis or for personal use, for example, to gather information about road traffic in the neighborhood or to present a quantitative evaluation of the traffic flow when selling a residence. To the best of our knowledge, such application does not exist. In the next section, we review the proposed system That makes the algorithm faster, requiring less memory space, minimizing the energy consumption, and real time on a smartphone. Let us detail the proposed counting approach.

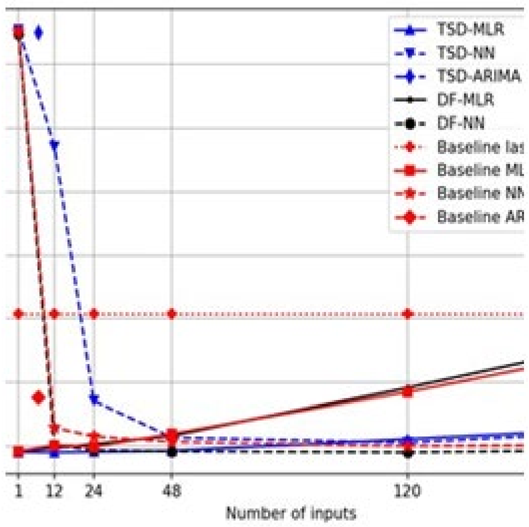

Fig 4 : Traffic Input Graph

CONCLUSION :

In this paper we have tackled the problem of forecasting a high frequency time series of mobile traffic (sampled every 5- minutes) starting from a lower-frequency time series of hourly aggregated traffic samples. We show through experiments that forecasting traffic sampled at different frequencies is possible and, especially for those cases where Neural Networks are used, the results even outperform forecasts obtained starting directly from high-frequency time series. This surprising result can be used as a starting point for several scenarios connected with cognitive networking and C-RAN optimizations (e.g., resource provisioning, storage optimization). Future works will better analyze the performance of other structures of Neural Networks when applied to the forecasting problem. In particular, we plan to add as input to the Neural Networks not only traffic samples but also other sources of data (e.g., number of users connected to the network or number of active sessions). Also, a fine-tuning of the Neural Network hyper-parameters may be performed, in order to increase the overall performance. Recent Neural Network architectures tailored to time series forecasting such as LSTM or RNN may also be tested.

REFERENCES

[1] J. Wu, Z. Zhang, Y. Hong, and Y. Wen, “Cloud radio access network (c-ran): a primer,” IEEE Network, vol. 29, no. 1, pp. 35–41, 2015.

[2] A. Jindal, V. Podolskiy, and M. Gerndt, “Multilayered cloud applications auto scaling performance estimation,” in 2017 IEEE 7th International Symposium on Cloud and Service Computing (SC2). IEEE, 2017, pp. 24–31.

[3] X. Zhou, Z. Zhao, R. Li, Y. Zhou, and H. Zhang, “The predictability of cellular networks traffic,” in 2012 International Symposium on Communications and Information Technologies (ISCIT). IEEE, 2012, pp. 973–978.

[4] D. Zhou, Z. Yan, Y. Fu, and Z. Yao, “A survey on network data collection,” Journal of Network and Computer Applications, vol. 116, pp. 9–23, 2018.

[5] G. C. Chow, A.-l. Lin et al., Best linear unbiased interpolation, distribution, and extrapolation of time series by related series. Princeton University., 1971.

[6] F. T. Denton, "Adjustment of monthly or quarterly series to annual totals: an approach based on quadratic minimization," Journal of the American Statistical Association, vol. 66, no. 333, pp. 99–102, 1971.

[7] J. Casals, M. Jerez, and S. Sotoca, "Modelling and forecasting time series sampled at different frequencies," Journal of Forecasting, vol. 28, no. 4, pp. 316–342, 2009.

[8] H. Wheater, V. Isham, D. Cox, R. Chandler, A. Kakou, P. Northrop, L. Oh, C. Onof, and I. Rodriguez-Iturbe, "Spatial-temporal rainfall fields: modelling and statistical aspects," Hydrology and Earth System Sciences Discussions, vol. 4, no. 4, pp. 581–601, 2000.

[9] R. R. Andrawis, A. F. Atiya, and H. El-Shishiny, "Combination of long term and short term forecasts, with application to tourism demand forecasting," International Journal of Forecasting, vol. 27, no. 3, pp. 870–886, 2011.

[10] 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 Communications Surveys & Tutorials, vol. 19, no. 3, pp. 1790– 1821, 2017.