Wireless Channel Quality Prediction using Sparse Gaussian Conditional Random Fields

[1]Raushan Raj, [2]Adita Kulkarni, [1]Anand Seetharam, [1]Arti Ramesh
[1]Department of Computer Science, SUNY Binghamton, [2]Department of Computing Sciences, SUNY Brockport
akulkarni@brockport.edu, {rrausha1, aseethar, artir}@binghamton.edu

Abstract*—Accurate wireless channel quality prediction over 4G LTE networks continues to be an important problem as future channel predictions are widely leveraged to meet the strict requirements of applications such as 360-degree video, AR/VR, and online games. The availability of large amounts of wireless channel data, the increase in computational power and the advancements in the field of machine learning provide us the opportunity to design learning-based approaches to address the channel quality prediction problem. In this paper, we design discriminative sequence-to-sequence probabilistic graphical models, specifically sparse Gaussian Conditional Random Fields (GCRF) models to accurately predict future channel quality variations in 4G LTE networks based on past channel quality data. In contrast to prior work that has primarily focused on designing parsimonious Markovian models or computationally-intensive deep learning models, the sparse GCRF models designed here provide superior performance while being highly interpretable and computationally efficient, thus making them an ideal choice for practical deployment. To validate the efficacy of our sparse GCRF model, we compare its performance (i.e., root mean squared error and mean absolute error) with *i)* linear regression and *ii)* ARIMA and *iii)* the state-of-the-art deep learning model on real-world 4G LTE channel quality data collected under varying levels of user mobility for two cellular operators and observe that the GCRF model provides significantly higher performance improvement.

***Index Terms*—channel quality estimation, Gaussian Conditional Random Fields**

I. Introduction

With the projected explosive increase in mobile wireless traffic in next generation cellular networks [1], accurately modeling and predicting wireless channel quality variations has become an extremely important problem. While a significant amount of research effort, starting from the seminal two-state Markov channel model [2] to the state-of-the-art deep learning models [3], has been devoted to predicting future channel variations, revisiting the channel quality prediction problem in cellular networks is critical for meeting the requirements of today's high bandwidth, low latency applications such as AR/VR, 360-degree videos and online games. The availability of computational power and large amount of 4G LTE channel quality data, together with the growth of machine learning algorithms present us the opportunity to design models that provide superior channel prediction, and to leverage these predictions for improving user quality of experience (QoE) and application-level performance over cellular networks.

In this paper, our goal is to design discriminative probabilistic graphical models, specifically spare Gaussian Conditional Random Fields (GCRF) models that achieve superior prediction performance of wireless channel variations (e.g., signal strength, uplink and downlink bitrate) in cellular networks. In comparison to existing analytical approaches that focus on designing Markovian models that exploit the underlying characteristics of the wireless medium (e.g., Rayleigh or Rician fading channels) [4], [5], and deep learning models [3], [6] for wireless channel modeling and prediction, we design spare GCRF models for this prediction task.

Sparse GCRF models have the unique characteristic that they provide good explainability and interpretability (similar to Markovian models) along with superior modeling and predictive power (similar to deep models), which make them an ideal candidate for the channel prediction problem. Additionally, the sparse nature of the model ensures that it takes requires only limited computational resources to train. For example, sparse GCRF models take significantly less time to train (i.e., only few minutes) when compared to deep models that usually take few hours. These attributes make sparse GCRFs an attractive model for practical deployment.

In order to make effective predictions, we cast wireless channel quality prediction as a time series prediction problem where the objective is to predict future variations (e.g., signal strength) based on past data. We then design spare GCRF models [7], a structured regression approach to successfully encode dependencies between past and future channel quality data. Sparse GCRF models take entire past sequences into account to make future predictions and thus generally outperform analytical and statistical models [8], [9]. As sparse GCRFs are probabilistic graphical models that only learn the required dependencies among the output variables and those between the input and the output, they are inherently interpretable.

To demonstrate the efficacy of our sparse GCRF model, we conduct experiments on channel quality data collected over 4G LTE networks in multiple vehicular (i.e., bus, car and train) and pedestrian mobility scenarios for two different operators. The dataset contains four channel quality metrics, namely RSRP, RSRQ, downlink and uplink bit rate collected simultaneously at a one second time granularity. We design separate models to predict each of these metrics and compare the performance of our model with the state-of-the-art LSTM based deep learning model and two baselines—linear regression and ARIMA. We observe that our model significantly outperforms the baselines in all scenarios, yielding a performance improvement of 20% on average. Our main performance evaluation metrics are the root mean squared error (RMSE) and the mean absolute error (MAE).

Overall, our analysis and results demonstrate that the trained GCRF model has limited data and computational requirements at test time, is highly interpretable and provides superior prediction performance. These qualities make it a practically beneficial model that can be deployed to improve application-level performance in 4G LTE networks.

II. Related work

Wireless channel quality prediction continues to be one of the most important problems in communications and networking research as an ever-increasing number of applications (e.g., AR/VR, IoT applications, online games) rely on good channel prediction for efficient data transfer and for achieving good user QoE. The earliest work in this space is the two state Gilbert and Elliot Markov model. Since then, a large number of Markovian models have been designed to provide good performance for a variety of different settings, environments and technology standards [10], [11].

Though Markovian models enhance our understanding of the wireless channel and help us appreciate the impact of various factors such as multi-path fading, shadowing, and path loss on performance, they usually suffer from poor prediction performance in real-world settings as they only take a limited amount of history into consideration while making these predictions. Therefore, in recent years, a number of statistical techniques (e.g., regression, filtering) and machine learning models, in particular deep learning models have been proposed for solving different problems in the communications domain. For example, deep learning models have been shown to provide good performance for geo spatio-temporal modeling and prediction in cellular networks [6], [12]–[16]. Similarly, in our recent work [3], we designed a deep learning model that provides good signal strength prediction in a variety of different settings and networks.

One of the main limitations of existing work that use machine learning techniques to solve problems in the communications and networking field is that majority of them use machine learning models as a black-box and limited attention has been paid to choosing the most appropriate model for a specific problem. Additionally, most research efforts in this space has relied on easy ready-to-use packages such as TensorFlow [17] and scikit-learn [18] and has seldom explored probabilistic models such as Conditional Random Fields or Markov Random Fields and their Gaussian and sparse variants that are usually available in the form of 'research' code. Probabilistic graphical models, particularly the sparse GCRF models that we explore here have been shown to provide great performance for a diverse set of prediction tasks [8], [9]. In contrast to prior work, we design GCRF models for the wireless channel quality prediction problem and aim to demonstrate the utility of this relatively unexplored, yet powerful model to address important problems in the communications domain.

III. Problem statement & GCRF prediction model

In this paper, our goal is to design models that can accurately model and predict wireless channel fluctuations in 4G LTE networks. In addition to maintaining traditional QoE guarantees, many applications such as 360-degree videos, AR/VR and online games depend strongly on accurate prediction of cellular network channel variations. In the next subsections, we present the problem statement and describe our sparse GCRF model for the channel quality prediction problem. We then validate the efficacy of our sparse GCRF model via experiments on real-world channel quality measurement data collected over multiple 4G LTE networks (Sections IV and V).

A. Problem Statement

We formulate the channel quality prediction problem as a time-series prediction problem where at each time T, one has a sequence of past measurement values in a window size of n (i.e., $X_T = [x_{T-n}, x_{T-(n-1)}, ..., x_{T-1}, x_T]$) that is used to predict variations for k steps into the future (i.e., $\hat{Y}_T = [\hat{y}_{T+1}, \hat{y}_{T+2}, ..., \hat{y}_{T+(k-1)}, \hat{y}_{T+k}]$). Note that Y_T denotes the actual measurement values. In this work, we consider and predict multiple different metrics of the wireless cellular channel, namely Reference Signal Received Power (RSRP), Reference Signal Received Quality (RSRQ), uplink bitrate, and downlink bitrate.

B. Why sparse GCRF models for channel prediction?

Before describing our solution, we present a case outlining the benefits of investigating GCRF models for solving various prediction tasks in the communications domain, an example of which is the channel quality prediction task studied here. Statistical techniques (e.g., ARIMA, regression) and Markovian models are highly interpretable and help in understanding the influence of the underlying characteristics (e.g., multi-path fading, path loss) on the quality of the wireless channel, but often provide below par prediction performance as they take limited amount of historical data into consideration and fail to capture the complex interdependencies in the data. To address performance issues, over the past few years, deep learning models have emerged as an attractive alternative, but they suffer from high computational requirements and lack of interpretability.

This provides us the opportunity to explore and develop discriminative sequence-to-sequence probabilistic graphical models, specifically sparse GCRF models that capture the best of both worlds (i.e., statistical models and deep learning models). Being graphical models, spare GCRF models are inherently interpretable as one can peruse through the learned weights and understand the connections responsible for generating the predictions. Additionally, being sequence-to-sequence models, sparse GCRF are ideal for time series prediction tasks that require mapping input sequences to output sequences. As sequence-to-sequence models take entire input sequences into consideration when predicting the future, sparse GCRFs provide superior performance than traditional analytical models and often in the same ballpark as deep learning models.

Another attractive attribute of sparse GCRF models is that they are computationally efficient and require only a fraction of the time needed to train deep learning models. These qualities make GCRF models highly desirable and practically feasible for many real-world applications [8], [9].

C. GCRF Prediction Model

GCRF models are the Gaussian variant of Conditional Random Fields (CRF) models [19]. CRFs are discriminative probabilistic graphical models that capture the dependencies between the output variables and those between the input and the output via a graphical model and do not explicitly encode the dependencies among the inputs. Thus, CRFs primarily focus on the relationship between the output variables, thus modeling the posterior distribution of the future wireless channel quality given the historical wireless channel quality. As channel quality prediction is a continuous-valued prediction problem and the basic CRF model is designed for discrete valued data [19], we consider a recent version of CRFs extended to structured regression, sparse GCRFs [7], for predicting future wireless channel quality variations. The sparse GCRF model that we design only captures the desired dependencies and thus informs us about the relationships in the graphical model that are useful for the prediction task.

The distribution modeled by GCRFs is given by,

$$P\left(Y_T \mid X_T; \Lambda, \Theta\right) = \frac{1}{Z(X_T)} exp(-Y_T' \Lambda Y_T - 2X_T' \Theta Y_T) \tag{1}$$

where Θ and Λ are parameters/regression coefficients of the GCRF model. Θ is the adjacency matrix containing the edges between x_t and y_t, while Λ is the inverse covariance matrix, containing the edges amongst the y_t's. The CRF is a Gaussian distribution with mean $-\Lambda^{-1}\Theta' X_T$ and variance Λ^{-1}. $Z(X_T)$ in Equation 1 is the partition function, which ensures that the posterior is integrated to 1.

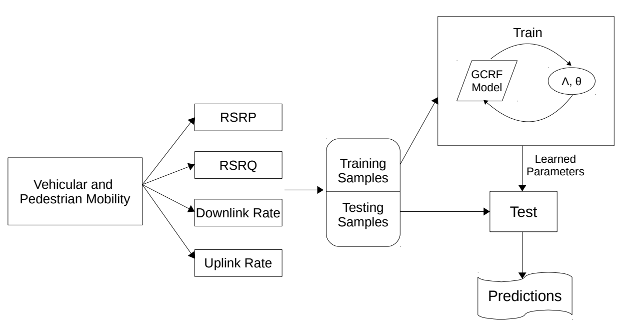


Fig. 1: GCRF Prediction Model Architecture

Figure 1 provides an overview of the architecture of the GCRF prediction model. The architecture comprises of a data pre-processing component, which pre-processes the wireless channel quality measurement data for the various channel metrics and the different mobility scenarios. We then split the processed data into train and test samples. The GCRF model then learns the values of the parameters based on the training data. Once the training process is complete, we have the trained GCRF model that uses the test samples and generates the desired predictions.

D. GCRF training and testing

At training time, the goal is to learn the values of the parameters Θ and Λ based on the training data. We use the maximum likelihood estimation approach to determine the parameters by maximizing the probability of the data given the parameters (i.e., the likelihood function).

$$\max_{\Lambda,\Theta} P\left(y \mid x; \Lambda, \Theta\right) \tag{2}$$

Maximizing the likelihood function in Equation 2 is equivalent to minimizing the negative log-likelihood, which is what we perform in our implementation (Equation 3).

$$\min_{\Lambda,\Theta} -log(P\left(y \mid x; \Lambda, \Theta\right)) \tag{3}$$

To avoid overfitting the model and to ensure that it generalizes better, we use regularization, an approach that penalizes high-valued regression coefficients. This is achieved by adding a regularization term to the maximum likelihood estimate equations above. As our objective is to learn a sparser set of regression coefficients by driving the less contributing coefficients to zero (i.e., the most desired dependencies), we use L_1 regularization. After expanding $P\left(y \mid x; \Lambda, \Theta\right)$, and adding L_1 regularization, the optimization function is as follows:

$$\min_{\Lambda,\Theta} -log|\Lambda| + \frac{1}{k} tr(Y^T Y \Lambda + 2Y^T X \Theta + \Lambda^{-1} \Theta^T X^T X \Theta \Lambda^{-1}) + \lambda(||\Lambda|| + ||\Theta||) \tag{4}$$

where k is the number of samples in X and Y. The last term $\lambda(||\Lambda|| + ||\Theta||)$ gives the regularization term, where λ is the regularization constant. We adopt the optimization method proposed by Wytock et al. [7] for solving the GCRF with L_1 regularization. This optimization approach uses a second-order active set method, which at first iteratively produces a second-order approximation to the objective function without the L_1 regularization term. The approximated objective function is then solved with L_1 regularization term using alternating Newton coordinate descent.

Figure 2 shows an example of a trained GCRF model that uses n input sample points to predict k steps into the future. We note that we start with a fully-connected graph between x and y and between y's. The use of L_1 regularization drives some of these edge values to zero, thus yielding a sparser graph after training. In Figure 2, we have omitted some edges between x and y to illustrate the sparser nature of the graphical model learned after training.

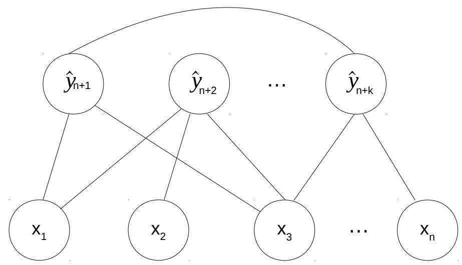


Fig. 2: Trained GCRF Model

E. GCRF Implementation Details

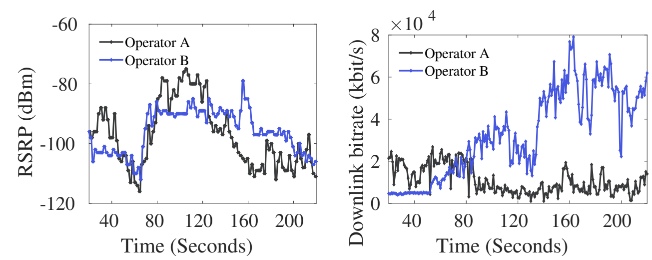
We use *SGCRFPy*, a Python toolkit for sparse GCRFs[1] to implement our models. The first 75% of the dataset is used for training the model and the remaining 25% is used for testing. At training time, our GCRF models use past n seconds as input to predict the next k seconds. We use a sliding window approach that moves the window one second at a time to cover the entire data. We use $n = 20$ and $k = 10$ in our experiments. The parameter Λ is initialized to the identity matrix and Θ is initialized to all zeros. Using regularization constant $\lambda = 0.1$ and $10,000$ iterations, the model converges on a set of dependencies learned from the training data. We train separate GCRF models for the wireless channel quality metrics for each mobility scenario for both network operators.

At test time, prediction for the next k seconds is generated from a test sequence of length n using the parameter values Λ and Θ learned at training time. The predicted values $\hat{Y}$ are compared with the ground truth channel quality values Y, and the RMSE and MAE metrics are computed. For one configuration of parameter values, the training time for the model is around 20 minutes on a standalone lab machine, while the testing time is only a couple of minutes. The memory requirement for executing the GCRF model is also low.

IV. Datasets and Data Preprocessing

We use a recently collected 4G LTE dataset [20] consisting of various channel quality metrics collected for two major mobile network operators (anonymized as Operator A and B) in Cork City, Ireland, in vehicular and pedestrian mobility scenarios using a Samsung J5 mobile phone. The G-NetTrack Pro mobile network monitoring tool is used for collecting the channel metrics. The vehicular measurements are collected by traveling in a **car**, **bus**, and **train**, whereas the **pedestrian** measurements are collected by walking around the city. The measurements are collected at one second granularity. Multiple traces are collected in every scenario and each trace is around 15 minutes long. We combine these traces to obtain a single one hour long trace for each mobility scenario for both network operators. The dataset contains around 2% missing values. We fill them using linear regression. The dataset consists of the following channel quality metrics, which we use for prediction.

- **RSRP**: Reference Signal Received Power represents the average power of LTE reference signals. It is measured in dBm.
- **RSRQ**: Reference Signal Received Quality represents the ratio between RSRP and Received Signal Strength Indicator (RSSI). It indicates the quality of the received reference signal and is measured in dBm.
- **Downlink bitrate**: It is the download rate measured at the device in kbit/s.
- **Uplink bitrate**: It is the uplink rate measured at the device in kbit/s.



(a) RSRP (Train) (b) Downlink bitrate (Pedestrian)

Fig. 3: Channel Quality Variations

Figure 3 shows the RSRP and downlink bitrate for the train and pedestrian mobility scenarios for both mobile operators, respectively. We observe from the figure that the channel quality metrics vary significantly over time for both operators. We hypothesize environmental factors, mobility, network coverage along with the operators' internal traffic policy as the primary reasons behind these variations [20].

V. Experimental Evaluation

In this section, we compare the performance of our GCRF model with the following models.

Linear regression: This model uses linear extrapolation for generating future predictions based on the input data.

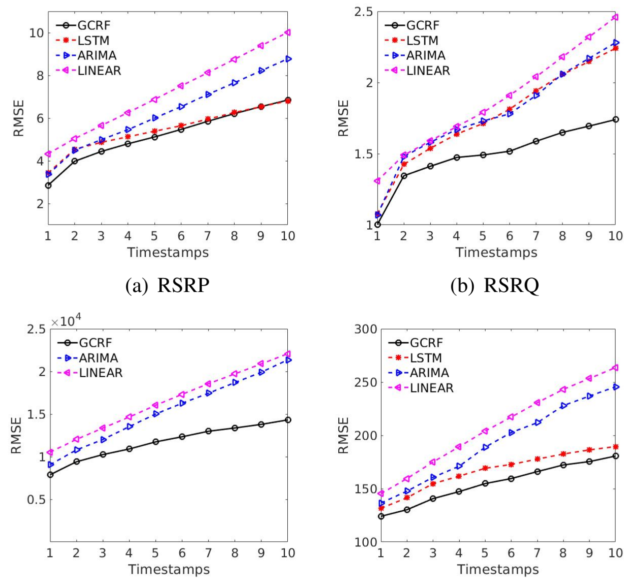
ARIMA (p, d, q): This model has three components—AR (autoregressive term), I (differencing term) and MA (moving average term), which are specified by p, d and q respectively. The number of past values used for predicting the future is denoted by p, the degree of differencing (i.e., the number of times the differencing operation is performed to make a series stationary) is denoted by d, and the number of error terms used to predict the future values is denoted by q.

LSTM based Deep Learning Model: This is the state-of-the-art LSTM based deep learning model for channel prediction [3]. It is a sequence-to-sequence model that uses recurrent neural networks as the underlying neural network architecture.

The main evaluation metrics used in our experiments are the root mean squared error (RMSE) and the mean absolute error (MAE), which are given by equations 5 and 6 respectively.

$$RMSE_j = \sqrt{\frac{\sum_{i=1}^{h} (\hat{y}_{ij} - y_{ij})^2}{h}} \tag{5}$$

[1]Sparse GCRF: https://github.com/dswah/sgcrfpy.



(a) RSRP (b) RSRQ

(c) Downlink bitrate (d) Uplink bitrate

Fig. 4: RMSE Results (Bus, Operator A)

$$MAE_j = \frac{\sum_{i=1}^{h} |\hat{y}_{ij} - y_{ij}|}{h} \quad (6)$$

where y_{ij} is the i^{th} test sample for j^{th} prediction step, $\hat{y}_{ij}$ is the predicted value of y_{ij}, and h is the total number of test samples.

We first present the RMSE and MAE results and then discuss the qualitative results. Due to lack of space, we mainly present results for the bus mobility scenario for operator A, though we obtain similar results for all other settings as well.

A. RMSE Results

In this subsection, we compare the RMSE performance of the sparse GCRF model with the baselines. Figure 4 shows the performance of the models for the channel quality measurements conducted on a bus for the different metrics (i.e., RSRP, RSRQ, downlink bitrate and uplink bitrate) for operator A. We observe from the figure that the deep model outperforms ARIMA and linear regression for RSRP and uplink, and provides similar performance to ARIMA for RSRQ. For downlink, the performance of the deep learning model is poor and hence we omit it from Figure 4(c). The primary reason is that variation in the downlink dataset is high, which throws off the prediction performance of the deep model. Additionally, we note that the datasets studied in this paper are different from the ones studied in [3]. In comparison, the GCRF model significantly outperforms the linear regression, ARIMA and the deep models for all channel quality metrics.

We also see that the RMSE values for the baselines increase considerably for each predicted time step further into the future, while there is a gradual increase in the RMSE values for the GCRF model. This shows that the GCRF model is able to predict future values more accurately than the baselines. We attribute this performance improvement to the sequence-to-sequence modeling behavior and the sparse nature of the GCRF model that enables it to learn only the most important dependencies.

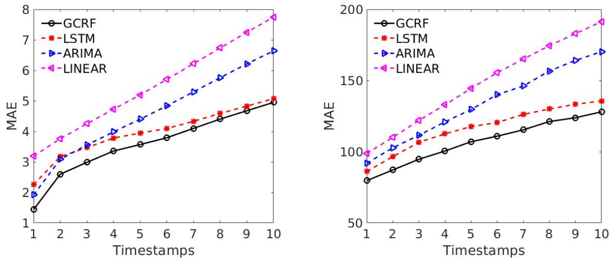
Table I shows the RMSE values for all the channel quality metrics calculated as an average over 10 future predicted time steps for all mobility scenarios and both network operators. We observe that GCRF outperforms the baselines for all channel quality metrics. The average performance improvement of GCRF over linear regression, ARIMA and LSTM is 24%, 15% and 16%, respectively.

TABLE I: Average RMSE Results

Vehicle Operator	Metrics	GCRF	LSTM	ARIMA	Linear
Bus Operator A	RSRP	**5.21**	5.46	6.26	7.2
	RSRQ	**1.49**	1.76	1.77	1.88
	DL	**11700.12**	20075.16	15411.92	16519.71
	UL	**155.1**	166.76	193.0	208.13
Bus Operator B	RSRP	**4.85**	5.2	5.53	6.01
	RSRQ	**1.21**	1.32	1.43	1.64
	DL	**8976.44**	13205.36	10763.86	11699.42
	UL	**180.49**	188.05	223.47	237.88
Car Operator A	RSRP	**5.11**	5.15	6.1	6.56
	RSRQ	**1.98**	2.46	2.28	2.52
	DL	**9332.36**	21885.73	10113.7	11062.18
	UL	**147.59**	147.79	157.26	166.3
Car Operator B	RSRP	**5.56**	6.05	7.04	7.66
	RSRQ	**1.23**	1.29	1.39	1.62
	DL	**13217.17**	29096.05	16826.13	19611.85
	UL	**243.63**	249.56	319.87	340.71
Pedestrian Operator A	RSRP	**3.15**	3.19	3.76	4.15
	RSRQ	**1.5**	1.51	1.69	1.94
	DL	**4383.06**	8461.25	5005.38	5350.42
	UL	89.18	**88.8**	106.59	107.85
Pedestrian Operator B	RSRP	4.06	**4.02**	4.91	5.08
	RSRQ	**1.34**	1.41	1.52	1.66
	DL	**10113.42**	21826.29	10802.04	12314.23
	UL	**185.4**	188.45	202.62	242.01
Train Operator A	RSRP	**7.69**	8.01	8.65	9.51
	RSRQ	**1.99**	2.0	2.37	2.72
	DL	**2977.97**	5710.54	3452.03	3670.14
	UL	**61.66**	64.35	66.47	75.16
Train Operator B	RSRP	**4.42**	**4.42**	5.45	6.53
	RSRQ	**1.29**	1.31	1.4	1.69
	DL	**1609.09**	2128.48	1903.57	2094.66
	UL	**46.77**	49.51	53.07	56.51

B. MAE Results

In this subsection, we discuss the performance of the models with respect to MAE. Due to lack of space we only present limited results for MAE. Figure 5 shows the performance of the models for RSRP and uplink for the bus mobility scenario for operator A while Table II shows the MAE values for all the channel quality metrics calculated as an average over 10 future predicted time steps considering all mobility scenarios and both network operators. Similar to the RMSE results, we observe that the GCRF model significantly outperforms the linear regression and ARIMA baselines for all channel quality metrics for MAE as well. GCRF shows 26%, 16% and 20% average improvement over linear regression, ARIMA and LSTM, respectively.



(a) RSRP (b) Uplink bitrate

Fig. 5: MAE Results (Bus, Operator A)

TABLE II: Average MAE

Metrics	GCRF	LSTM	ARIMA	Linear
RSRP	**3.37**	3.68	4.13	4.7
RSRQ	**0.94**	1.05	1.07	1.26
DL	**5159.08**	10745.31	6124.48	6958.74
UL	**86.39**	93.64	103.8	115.9

C. Qualitative Results

In this subsection, we discuss the qualitative performance of the models by comparing their one step predictions with the actual values. Due to lack of space, we present representative results for RSRP for the bus mobility scenario for operator A. Figure 6 shows the predictions of the GCRF model, ARIMA and LSTM. From the figure, we observe that the predictions generated by the GCRF model match the actual values more closely in comparison to ARIMA and LSTM. This is because GCRF takes entire input sequences in account to predict the future and is thus able to make more informed decisions in comparison to the statistical techniques that attempt to model the data using linear extrapolations.

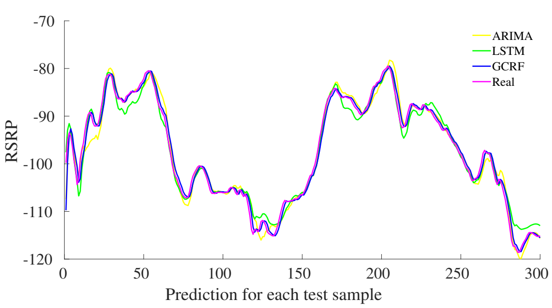


Fig. 6: Qualitative Results: GCRF vs ARIMA

VI. Conclusion

In this paper, we investigated the channel quality prediction problem in 4G LTE networks. We developed discriminative probabilistic graphical models, specifically sparse GCRF models to accurately predict future wireless channel quality variations (i.e., RSRP, RSRQ, uplink bitrate, downlink bitrate) by taking historical channel quality data into account. We compared the performance (i.e., RMSE, MAE) of our GCRF model with the state-of-the-art deep model and the linear regression and ARIMA baselines on real-world 4G LTE data collected over multiple operators and observed that our model significantly outperforms these baselines. Our experiments demonstrate that the GCRF model is computationally efficient and highly interpretable, thus making it practically feasible.

References

[1] Ericsson report on mobile data traffic growth outlook). https://www.ericsson.com/en/mobility-report/reports/november-2018/mobile-data-traffic-growth-outlook.

[2] Edgar N Gilbert. Capacity of a burst-noise channel. *Bell system technical journal*, 39(5):1253–1265, 1960.

[3] Adita Kulkarni, Anand Seetharam, Arti Ramesh, and J Dinal Herath. Deepchannel: Wireless channel quality prediction using deep learning. *IEEE Transactions on Vehicular Technology*, 69(1):443–456, 2019.

[4] Claude Oestges, Nicolai. Czink, Bernd Bandemer, and Paolo Castigoline. Experimental characterization and modeling of outdoor-to-outdoor and indoor-to-indoor distributed channels. *IEEE Trans. On Vehicular Technology*, 59, 2010.

[5] Yu Zhang, Jianhua Zhang, Di Dong, Xin Nie, Guanngyi Liu, and Ping Zhang. A novel spatial autocorrelation model of shadow fading in urban macro environments. In *IEEE GLOBECOM*, 2008.

[6] Ayon Chakraborty, Md Shaifur Rahman, Himanshu Gupta, and Samir R Das. Specsense: Crowdsensing for efficient querying of spectrum occupancy. In *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*, pages 1–9. IEEE, 2017.

[7] Matt Wytock and Zico Kolter. Sparse gaussian conditional random fields: Algorithms, theory, and application to energy forecasting. In *International conference on machine learning*, pages 1265–1273, 2013.

[8] Gisella Bejarano, Adita Kulkarni, Raushan Raushan, Anand Seetharam, and Arti Ramesh. Swap: Probabilistic graphical and deep learning models for water consumption prediction. In *Buildsys*. EAI, 2019.

[9] David DeFazio, Arti Ramesh, and Anand Seetharam. Nycer: A non-emergency response predictor for nyc using sparse gaussian conditional random fields. In *on Mobile and Ubiquitous Systems: Computing, Networking (MobiQuitous), 2018 EAI International Conference on*. EAI, 2018.

[10] Michele Zorzi, Ramesh R Rao, and Laurence B Milstein. ARQ error control for fading mobile radio channels. *IEEE Transactions on Vehicular Technology*, 46(2):445–455, 1997.

[11] Anand Seetharam, Jim Kurose, and Dennis Goeckel. A markovian model for coarse-timescale channel variation in wireless networks. *IEEE Trans. Vehicular Technology*, 65(3):1701–1710, 2016.

[12] Jing Wang, Jian Tang, Zhiyuan Xu, Yanzhi Wang, Guoliang Xue, Xing Zhang, and Dejun Yang. Spatiotemporal modeling and prediction in cellular networks: A big data enabled deep learning approach. In *INFOCOM 2017-IEEE Conference on Computer Communications, IEEE*, pages 1–9. IEEE, 2017.

[13] Avik Ray, Supratim Deb, and Pantelis Monogioudis. Localization of lte measurement records with missing information. In *IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications*, pages 1–9. IEEE, 2016.

[14] Robert Margolies, Richard Becker, Simon Byers, Supratim Deb, Rittwik Jana, Simon Urbanek, and Chris Volinsky. Can you find me now? evaluation of network-based localization in a 4g lte network. In *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*, pages 1–9. IEEE, 2017.

[15] Rita Enami, Dinesh Rajan, and Joseph Camp. Raik: Regional analysis with geodata and crowdsourcing to infer key performance indicators. In *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6. IEEE, 2018.

[16] Emmanouil Alimpertis, Athina Markopoulou, Carter Butts, and Konstantinos Psounis. City-wide signal strength maps: Prediction with random forests. In *The World Wide Web Conference*, pages 2536–2542. ACM, 2019.

[17] Tensorflow. https://www.tensorflow.org/.

[18] Scikit learn. https://scikit-learn.org/stable/.

[19] John Lafferty, Andrew McCallum, and Fernando CN Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. 2001.

[20] Darijo Raca, Jason J Quinlan, Ahmed H Zahran, and Cormac J Sreenan. Beyond throughput: a 4g lte dataset with channel and context metrics. In *Proceedings of ACM Multimedia Systems Conference (MMSys 2018)*. Association for Computing Machinery, ACM, 2018.