Capacity Estimation of MIMO Systems via Support Vector Regression

Xiaolin Hu and Nicholas E. Buris
School of Communication and Information Engineering, Shanghai University, Shanghai, P.R. China
E-mail: {xiaolinhu, neburis}@shu.edu.cn

***Abstract*— The achievable capacity of Multi Input Multi Output (MIMO) capable wireless systems over a geographic area is an important quantity for system planning. Given the characteristics of the propagation environment and the specifics of the antenna systems on both sides of the link, the capacity can be evaluated in a straightforward and 'exact' albeit tedious and time consuming manner. This paper employs a Support Vector Regression (SVR) approach to create a machine learning model of the capacity and thus evaluate it fast, requiring only a fraction of 'exact' calculations. For purposes of system planning, it is seen that reasonable accuracy can be achieved with SVR at about 10-15% of the computational effort of the 'exact' approach.**

***Keywords*— MIMO, capacity, support vector regression, MEA, cartography.**

I. Introduction

MIMO architectures and techniques have become a standard feature in WiFi and cellular wireless systems. Not only can MIMO mitigate adverse effects of multipath, but it can also increase capacity via smart algorithms dictated by the system on the Multi Element Antennas (MEAs) at either side of a communication link. Given the precise characteristics of the MEAs involved as well as the precise propagation environment in the form of a superposition of plane waves departing from the transmitter and arriving at the receiver, the efficiency of a link under various smart algorithms can be calculated 'exactly'. For Single Input Single Output (SISO) systems this link efficiency was first defined and derived by Shannon [1], [2]. It was then extended by Foschini and Gans for the Open Loop transmission mode in [3]. Subsequent work outlined the link efficiency of the beamforming and the waterfilling algorithms [4]. The latter is the ultimate link efficiency under a constant available power constraint and both, the open loop and the beamforming algorithms, can be thought of as special cases of waterfilling.

The aforementioned discussion and the analyses in the present work are agnostic of the communications protocol used. Therefore, only capacity results, i.e. the maximum possible throughput, not the actual throughput, will be discussed. Moreover, capacity (in bps) is the link efficiency (in bps/Hz) multiplied by the available bandwidth for the transmitted signals. Because our analyses in this paper are independent of the frequency bandwidth, we will use the terms capacity and link efficiency interchangeably.

In order to plan the deployment of a wireless communications network, precise knowledge of the throughput that could be achieved in a geographic area is important and used to, among other things, determine the density of the base stations required for adequate coverage. For this throughput, the network planners have to take into account the number of users, their individual throughput requirements based on the services and applications that they run, the Quality of Service (QoS) in terms of the acceptable Bit Error rate (BER), the available Modulation and Coding Schemes (MCS), etc. It is not the intension of the present work to cover such network planning. Instead, the present work outlines a machine learning (ML) assisted method of arriving at efficient estimations of the theoretically maximum throughput possible in a geographical area of interest. Other researchers have used ML for cartography [5], [6], [7]. Those methods are focused on SISO wireless systems and just the RF power in the area of interest. More recent work has applied deep learning to other aspects of MIMO systems, such as channel state information (CSI) recovery and direction-of-arrival (DOA) estimation [8], [9]. However, deep learning requires a large amount of training data, which is limited in cartography tasks. The objective of the work in this paper is to find the characteristics of SVR created models for capacity estimations of arbitrary MIMO systems.

II. Problem Set-Up

We set up the following problem. We randomly distribute 2^{10} Rx radios in a 200 m x 200 m geographical area, $\mathcal{R} \subset \mathbb{R}^2$. We also employ two Tx radios. One transmits the desired signal; the other transmits interference. Let $f(\boldsymbol{x})$ denote the channel capacity between the Tx radio and the Rx located at $\boldsymbol{x}$. All radios can be equipped by MEA systems. Given the plane wave decomposition of the propagation environment in the space of interest, we have a specific collection of plane waves connecting each Tx-Rx radio pair. For the purposes of analyzing just capacity, the source of the Tx is modeled by a Thevenin equivalent Voltage source and its associated source impedance. The Rx is in turn modeled by the load impedance seen by the Rx antenna. The model for this Tx-Rx is shown in Fig. 1. Using an electromagnetics exact formulation including, among other things, the terminations of the antenna elements involved, we derive the transfer function from the Tx voltage source to the received voltages at each antenna load. These voltage transfer functions are directly related to the Antenna to Antenna (A2A) channel matrices. The latter are typically agnostic of the impedances involved by virtue of various, often unrealistic, oversimplifications of the antenna systems.

978-1-7281-3516-8/19/$31.00

The A2A transfer functions are then used to obtain the Open Loop link efficiency (capacity per frequency bandwidth) of each Tx-Rx link. We first obtain the capacity at all the Tx-Rx pairs in an exact manner using MIMObit® [10]. Then, we use the capacity of a subset of Tx-Rx pairs as “training data” in a machine learning scheme and derive a “ML capacity model” which is used to estimate the capacity everywhere, not just at the location of the training Rx radios. We then compare the capacity estimated by the ML capacity model to the ‘exact’ capacity as calculated by MIMObit.

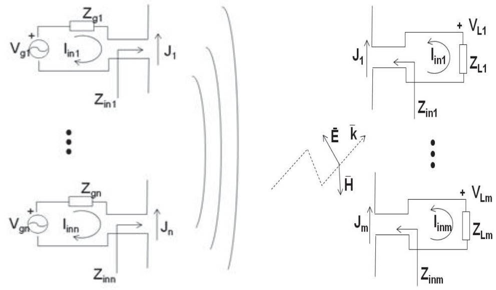

Fig. 1 Tx-Rx pair model. (a) N-port MEA system (Tx) and Thevenin equivalent sources. (b) M-port MEA system (Rx) and its termination loads.

III. Methods for Capacity Estimation

A. Support Vector Regression

This section briefly reviews the theory of SVR-based learning necessary to estimate capacity, following prior published similar work [5]. Given a set of training pairs $\{(\boldsymbol{x}_i, y_i)\}_{i=1}^{N}$, where N is the number of the training samples, $\boldsymbol{x}_i \subset \mathbb{R}^2$ is the i-th training samples and $y_i \in \mathcal{R}$ is the target value corresponding to $\boldsymbol{x}_i$. The goal of the regression problem is to find a function f that minimizes the loss function. In the SVR model, the target function is formulated as:

$$f(\boldsymbol{x}) = \phi(\boldsymbol{x})^T \boldsymbol{w} + b \tag{1}$$

where $\phi(\boldsymbol{x})$ denotes the nonlinear map function transforming $\mathbb{R}^2$ into high-dimensional feature space, $\boldsymbol{w}$ is the weight vector, b is a constant offset. By introducing $\epsilon -$ insensitive in the loss function [11], the optimization problem of SVR can be defined as follows:

$$min\left\{\frac{1}{2}\|\boldsymbol{w}\|^2 + C\sum_{i=1}^{N}(\xi_i + \hat{\xi}_i)\right\} \tag{2}$$
$$s.t\ (\phi(\boldsymbol{x_i})^T\boldsymbol{w} + b) - y_i \le \epsilon + \xi_i$$
$$y_i - (\phi(\boldsymbol{x_i})^T\boldsymbol{w} + b) \le \epsilon + \hat{\xi}_i$$
$$\epsilon \ge 0;\quad \xi_i, \hat{\xi}_i \ge 0,\quad i = 1, 2, \dots, N$$

The parameter ϵ denotes the allowed deviation between y_i and $\hat{f}(\boldsymbol{x_i})$. Slack variables ξ_i and $\hat{\xi}_i$ are used to measure the deviation lager than ϵ. The regularization parameter, $C > 0$, controls the tradeoff between the errors over the training data and the flatness of the target function [12]. We use the Lagrangian multipliers technique to solve the dual form of optimization problem (2), then the regression function is obtained:

$$\hat{f}(\boldsymbol{x}) = \sum_{i=1}^{N}(\hat{\alpha}_i - \alpha_i)\, k(\boldsymbol{x}, \boldsymbol{x}_i) + b \tag{3}$$

where $\hat{\alpha}_i$ and α_i are the Lagrangian multipliers, and $k(\boldsymbol{x}, \boldsymbol{x_i}) = \phi(\boldsymbol{x})^T\phi(\boldsymbol{x_i})$ is known as kernel function. In our work, LIBSVM [13] is used as the implementation of SVR and Gaussian function $\phi(\boldsymbol{x})^T\phi(\boldsymbol{x_i}) = \exp\left(-\gamma\|\boldsymbol{x} - \boldsymbol{x_i}\|^2\right)$ is selected as kernel function [14].

B. Gaussian Process Regression

Similarly to SVR, Gaussian process regression (GPR) is another kernel-based machine learning algorithm [15]. It is based on the assumption that the observed error of the target value $\boldsymbol{y}$ exhibits a standard normal distribution. GPR formulates a probabilistic model for solving regression problems within a Bayesian framework. Gaussian process is a distribution over functions. It is specified by a mean function and a kernel (covariance) function. Thus, the choice of kernel function is crucial for both GPR and SVR. Both techniques are tested in our experiments. The results indicate that SVR is superior to GPR on the capacity estimation task.

IV. Simulations And Results

All simulations in this paper are performed using a 4x2 MIMO scheme. Specifically, the signal Tx radio employs a 4-port MEA consisting of two pairs of ±45° halfwave dipoles λ/2 apart. The Rx antenna system consists of two halfwave vertical dipoles λ/2 apart. The “true” capacity in the area 200 m x 200 m is generated by solving the full problem using MIMObit and 2^{10} Rx radios. Then, the locations of sensors whose capacity is used as training data are selected randomly from those 2^{10} Rx radios. We evaluate the performance of the estimator versus interference power, number of sensors and propagation models. All analyses are carried out over WiFi frequency (2450MHz).

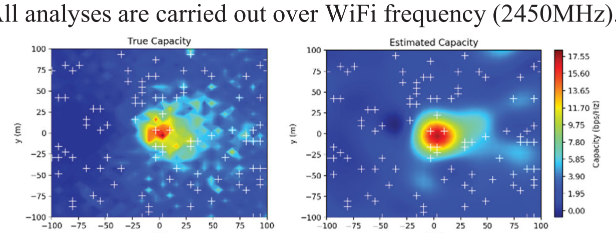

Fig. 2 True and SVR estimated capacity of a 4x2 MIMO WiFi system. The white crosses represent the locations of sensors.

Fig. 2 shows the “true” and estimated capacity over the interested area. The number of sensors is 90, about 9% of the Rx radios. The estimated capacity matches well the true one.

To measure the estimation performance, the normalized mean square error (NMSE) is defined as:

$$NMSE := \frac{E\left\{\left\|f(\boldsymbol{x}) - \hat{f}(\boldsymbol{x})\right\|^2\right\}}{E\{\|f(\boldsymbol{x})\|^2\}} \tag{4}$$

where $E[\cdot]$ denotes the expected value operation.

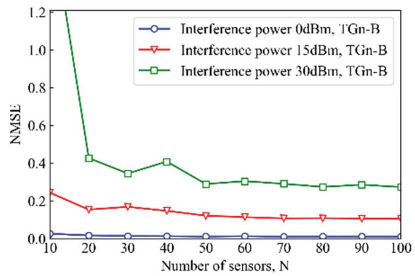

Fig. 3 SVR estimated capacity NMSE (three interference power values).

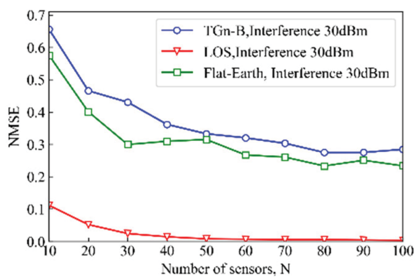

Fig. 4 SVR estimated capacity NMSE for different propagation models.

V. Discussion

Fig. 3 depicts the NMSE for different values of N and interference power. The propagation model is set as TGn-B. The total interference power is 0 dBm, 15 dBm and 30 dBm, respectively. It can be seen that the performance improves with N for different interference power values and the NMSE of lower interference has better performance. Here the reason is that higher interference resulted in more complex features, which are difficult to capture with SVR.

To illustrate the effects of scene complexity, Fig. 4 shows the NMSE as a function of N for different propagation models. Flat-Earth is a two-ray model and line-of-sight (LOS) corresponds to direct path propagation. Both of these models are simple and deterministic. While TGn-B is an indoor IEEE WiFi propagation model comprised of two random clusters of plane waves [16]. The first cluster has five plane waves, while the second has seven. The SVR runs better when the actual model is simple. As anticipated, the NMSEs are larger for TGn-B as compared to LOS and Flat-Earth. This is because the LOS and Flat-Earth models are deterministic and just a couple of parameters can capture their physics. However, in the TGn-B model, even though its major attributes are governed by given statistics, the specific instantiations are random and, thus, its physical behavior is not as easily captured.

VI. Conclusions

A study of capacity estimation in MIMO wireless system using the SVR algorithm is pursued. The performance of the estimator is tested from different aspects, including interference power, propagation model and the number of the training data. It is found that the cartography of capacity for such a MIMO link is feasible with an appreciable accuracy even when using just 10% of data for training. Although deterministic propagation environments are easy to model, empirical propagation models characterized by a degree of randomness are also adequately models with a small percentage of training data.

VII. References

[1] C. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal,* vol. 27, pp. 379-423, July 1948.

[2] C. Shannon, "Communication in the Presence of Noise," *Proceedings of the IRE,* vol. 37, no. 1, pp. 10-21, January 1949.

[3] G. Foscini and G. Gans, "On Limits of Wireless Communications in a Fading Environment when Using Multiple Antennas," *Wireless Personal Communications,* vol. 6, pp. 311-335, 1998.

[4] A. Goldsmith, Wireless Communications, Cambridge University Press, 2005.

[5] D. Romero, S.-J. Kim, G. B. Giannakis and R. Lopez-Varcarce, "Learning Power Spectrum Maps from Quantized Power Measurements," *IEEE Trans. Sig. Process.,* vol. 65, no. 10, pp. 2547-2560, 2017.

[6] M. Hamid and B. Beferull-Lozano, "Non-parametric spectrum cartography using adaptive radial basis functions," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, 2017.

[7] Y. Teganya, L. M. Lopez-Ramos, D. Romero and B. Beferull-Lozano, "Localization-Free Power Cartography," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, 2018.

[8] C.-K. Wen, W.-T. Shih and S. Jin, "Deep Learning for Massive MIMO CSI Feedback," *IEEE Wireless Commun. Lett.,* vol. 7, no. 5, p. 748–751, 2018.

[9] H. Huang, J. Yang, H. Huang, Y. Song and G. Gui, "Deep Learning for Super-Resolution Channel Estimation and DOA Estimation Based Massive MIMO System," *IEEE Trans. Veh. Technol.,* vol. 67, no. 9, p. 8549–8560, 2018.

[10] N. Buris. [Online]. Available: www.nebens.com.

[11] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and computing,* vol. 14, no. 3, pp. 199-222, 2004.

[12] J.-M. Bourinet, "Rare-event probability estimation with adaptive support vector regression surrogates," *Reliability Engineering & System Safety,* vol. 150, p. 210–221, 2016.

[13] C.-C. Chang and C.-J. Lin, "LIBSVM: A Library for Support Vector Machines," *ACM Trans. Intell. Syst. Technol.,* vol. 2, no. 3, pp. 1-27, 2011.

[14] M. Ring and B. M. Eskofier, "An approximation of the Gaussian RBF kernel for efficient classification with SVMs," *Pattern Recognition Letters,* vol. 84, p. 107–113, 2016.

[15] C. K. Williams and C. E. Rasmussen, Gaussian processes for machine learning, Cambridge: MIT press, 2006.

[16] V. Erceg et al., "IEEE Mentor," 10 May 2004. [Online]. Available: https://mentor.ieee.org/802.11/dcn/11/11-11-0968-04-00ah-channel-model-text.docx. [Accessed 28 6 2019].