A Machine Learning Approach to Model the Received Signal in Molecular Communications

H. Birkan Yilmaz, Changmin Lee, Yae Jee Cho, and Chan-Byoung Chae
School of Integrated Technology, Institute of Convergence Technology
Yonsei University, Korea.
Email:{birkan.yilmaz, cm.lee, yjenncho, cbchae}@yonsei.ac.kr

Abstract—A molecular communication channel is determined by the received signal. Received signal models form the basis for studies focused on modulation, receiver design, capacity, and coding depend on the received signal models. Therefore, it is crucial to model the number of received molecules until time \( t \) analytically. Modeling the diffusion-based molecular communication channel with the first-hitting process is an open issue for a spherical transmitter. In this paper, we utilize the artificial neural networks technique to model the received signal for a spherical transmitter and a perfectly absorbing receiver (i.e., first hitting process). The proposed technique may be utilized in other studies that assume a spherical transmitter instead of a point transmitter.

I. INTRODUCTION

The manipulation of matter at the atomic and the molecular scale constitutes nanotechnology. It is a promising technology that has numerous potential applications [1]–[3]. One of its innovative approaches is to utilize collaborative behavior amongst small entities. To enable the revolutionary possibilities of nanotechnology, it is important to possess the capacity to communicate at the nano- and micro-scale [4]. As a possible means to communication at such a small scale, researchers have proposed molecular communication via diffusion (MCvD) [5]. Examples of MCvD may be found prevalently in nature–quorum sensing between bacteria at micro-scale and pheromone-based communication between animals at macro-scale in both air and water environments. What occurs in these examples is molecules are utilized to convey information [1]–[3], [6]–[10]. In an MCvD system, a transmitter node emits molecules and molecules propagate through the environment until they arrive at the receiver node. Received molecules constitute the received signal. This is of prime importance when it comes to modeling an MCvD channel.

One of the main challenges in MC is to develop valid models for representing the received signal in different environments and conditions. Some of the MCvD models in the literature, assume that whenever a molecule hits the receiver it is removed from the environment [11]–[16]. This phenomena is modeled by the first-passage process (a.k.a first-hitting process). In this model, each molecule can contribute to the received signal only once. In [12], the authors presented the analytical model for the received signal in a 1-dimensional (1D) environment while considering the first-passage process. In [13], the authors enhanced the model from [12] by incorporating the drift component. In [15], the authors presented the expected cumulative number of received molecules when the transmitter is a point source and the receiver is an absorbing spherical node in a 3D medium. In [17], the authors compared the arrival process models used with the expected received signal that was derived in [15]. In [18], the authors analytically modeled the received signal and derived the expected cumulative number of received molecules when the receptor effect was added to the system presented in [15].

On the other hand, some of the MCvD models ignore the first-passage process, allowing molecules to pass through the receiver node surface/boundary with no interaction between the environment and the receiver [19]–[21]. In such models, the molecules are allowed to contribute to the signal multiple times, as they can pass in and out of the receiver node surface in and out multiple times.

Both of the MCvD physical layer models are summarized and presented in [4]. Researchers [4], [22] have claimed that the first-hitting process is observed in nature more frequently than the passive receiver process [4], [22]. Moreover, some of these models are validated by macro-scale testbed implementations [23]–[25]. Both channel models (i.e., the received signal for the first-hitting process and passive receiver) have a common hurdle for the communication engineering: heavy tail distribution of the received signal, which causes inter-symbol interference (ISI). In literature, plenty of techniques are proposed to eliminate the severe effects of ISI by utilizing specialized modulations [26]–[32], error correcting codes [33]–[35], and enzymes [19], [36]–[39]. In these studies, analytical models and their assumptions are used for the derivations, which indicates the importance of modeling the channel and the received signal in different environments.

Modeling the MCvD channel with the first-hitting process is an open issue for a spherical transmitter. The literature introduces analytical models for the first-hitting process with a point transmitter and with a spherical receiver [15]. In this paper, we propose a channel model that is inspired from the derived formulation in [15]. The proposed model function has model parameters that are affected by the system parameters. Moreover, we show that these model parameters can be learned by an artificial neural network (ANN) and the trained network can predict the model parameters for a given system parameters.
II. SYSTEM MODEL

In an MCvD system, there is at least one transmitter and receiver pair in a fluid environment. Figure 1 shows two different cases for MCvD—point and spherical transmitters. In general, the point transmitter case is studied in the literature [15], [18], [21]. The point transmitter assumption is reasonable for some applications. However, the transmitter node has a body in general and the transmitter does not react to the molecules from itself. In this paper, we model the MCvD channel with a perfectly absorbing receiver (i.e., whenever a molecule hits the surface of the receiver node, it contributes to the received signal) without the point transmitter assumption.

As shown in Fig. 1, emitted molecules diffuse in the 3D environment, which is characterized by the diffusion coefficient $D$. At the receiver side, the radius of the receiver is denoted by $r_{Rx}$ and the received signal consists of the time histogram of hitting molecules. When we have a point transmitter, molecules are able to move freely in the opposite direction of the receiver more freely. In the spherical transmitter case, however, molecules are obstructed and reflected by the transmitter of radius $r_{Tx}$. Hence, the received signals of these two cases are expected to differ.

A. MCvD with a Point Transmitter

The received signal is modeled analytically for the point transmitter case. The diffusion process basically models the average movement of particles in the concentration gradient. For modeling the received signal, we need to include the boundary conditions that represent the concentration and molecule distribution function. First, we consider Fick's Second Law in a 3D environment, which is given by

$$\frac{\partial p(r,t|r_0)}{\partial t} = D\nabla^2 p(r,t|r_0)$$

(1)

where $\nabla^2$, $p(r,t|r_0)$, and $D$ are the Laplacian operator, the molecule distribution function at time $t$ and distance $r$ given the initial distance $r_0$, and the diffusion constant. The initial condition is given by

$$p(r,t \to 0|r_0) = \frac{1}{4\pi r_0^2} \delta(r - r_0),$$

(2)

and the boundary conditions by

$$\lim_{r \to \infty} p(r,t|r_0) = 0,$$

(3)

$$D \frac{\partial p(r,t|r_0)}{\partial r} = w p(r,t|r_0), \text{ for } r = r_{Rx}$$

(4)

where $r_{Rx}$ and $w$ denote the radius of the receiver and the rate of reaction. The reaction rate with the receiver boundary is controlled by $w$. Specifically, $w = 0$ means a nonreactive surface and $w \to \infty$, on the other hand, corresponds to the boundary where every collision leads to an absorption.

In [15], the solution to this differential equation system is presented and analyzed from the perspective of channel characteristics. After finding the reaction rate, the authors presented the formula for the fraction of molecules that hit the receiver until time $t$, as follows:

$$F_{hit}^{3D}(t) = \frac{r_{Rx}}{d + r_{Rx}} \text{erfc} \left( \frac{d}{\sqrt{4Dt}} \right) = \frac{2r_{Rx}}{d + r_{Rx}} \text{erf} \left( \frac{d}{\sqrt{2Dt}} \right)$$

(5)

where $d$, erfc$(\cdot)$, and erf$(\cdot)$ represent the distance, complementary error function, and the standard Gaussian cdf, respectively. When the transmitter is a point, there is a circular symmetry, all the points at the same radius are equivalent and the solution for the system of differential equations is enabled. For the spherical transmitter case, however, it is more complex to derive the formulation of the number of received molecules.

B. MCvD with a Spherical Transmitter

Modeling the received signal analytically for a spherical transmitter is an open issue when the receiver is a perfectly absorbing spherical receiver in a 3D environment. The main difference and hurdle stem from the lack of circular symmetry. For the spherical transmitter case, molecules are biased towards going in the direction of the receiver due to the obstructing body of the transmitter node. Therefore, each of the molecules is expected to have a higher probability of hitting the receiver.

Hurdles caused by a spherical transmitter steered us to simulate the MCvD with a spherical transmitter and to analyze the patterns so as to grasp the underlying dynamics. We ran extensive simulations with many parameters. We realized that the cumulative number of received molecules exhibits a similar behavior to the point transmitter case with a small perturbation that is dependent on the system parameters.

III. PROPOSED TECHNIQUE FOR CHANNEL MODELING

As noted above, the main challenge in MCvD is to develop valid models for representing the received signal in different environments and conditions. When the transmitter has a spherical body, we propose to model the received signal by
The proposed technique is composed of two main phases: fitting the model parameters constitutes phase one (i.e., forming the input-output dataset for phase two) and learning the patterns in the input-output dataset constitutes phase two. After the learning phase (i.e., phase two), the output of the algorithm is a trained ANN for future predictions on unexplored input. A representative scheme of the proposed technique is depicted in Fig. 2.

### A. Model Function and Fitting

We propose two different model functions for fitting the simulation data and name them as **primitive model** and **enhanced model**. We use only a scaling factor for the primitive model, which is represented as follows:

\[
F_{hit}^{3D}(t, b_1) = b_1 \frac{r_{Rx}}{d + r_{Rx}} \text{erfc} \left( \frac{d}{\sqrt{4Dt}} \right)
\]

where \(b_1\) represents the model fitting parameter. For the enhanced model, we also parametrize the components related to \(D\) and \(t\), which is shown as follows:

\[
F_{hit}^{3D}(t, b_1, b_2, b_3) = b_1 \frac{r_{Rx}}{d + r_{Rx}} \text{erfc} \left( \frac{d}{(4D)^{b_2} t^{b_3}} \right)
\]

where \(b_1, b_2,\) and \(b_3\) are model fitting parameters. These model fitting parameters are introduced for fitting simulation data to model functions.

To find the model parameters corresponding to the system parameters, we use nonlinear least squares curve fitting technique. Assuming that we have \(N\) observations during the simulation we formulate the fitting model parameter estimation problem with \(m\) parameters as follows:

\[
\arg \min_{b_1, ..., b_m} \sum_{k=1}^{N} \left( F_{hit}^{3D}(t_k, b_1, ..., b_m) - S_{hit}^{3D}(t_k) \right)^2
\]

where \(S_{hit}^{3D}(t)\) corresponds to simulation data that is representing the ratio of hitting molecules until time \(t\).

For primitive and enhanced models, we apply to the simulation data nonlinear least squares curve-fitting technique. The output of the curve fitting process consists of the model parameters (i.e., \(b_1\) for primitive model and \(b_1 \sim b_3\) for the enhanced model). Hence, we obtain model parameters for each simulation case, which forms the dataset of the next phase. This dataset structure is depicted in Fig. 2 in the middle panel, which contains simulation case system parameters and the model parameters (i.e., \(b_i\)’s) from curve fitting.

### B. Learning Model Parameters

In this paper, we introduce a machine-learning technique to model the received signal in MCvD with a spherical transmitter. One of the popular machine-learning techniques is artificial neural networks. They have simple neuron-like nodes with thresholds and the connections with weights. Basically, the thresholds and the weights are adjusted until the desired output is observed for the given inputs.

The dataset from the curve fitting phase is divided into two disjointed subsets as training and validation datasets. Training data is utilized for training the ANN for getting the desired output for given inputs. Bayesian regularization backpropagation technique is used for the ANN training that updates the weights and bias values according to Levenberg-Marquardt optimization. Bayesian regularization minimizes a combination of squared errors and weights to determine the ANN parameters that generalize the pattern in the input-
output pairs. We utilize the trained ANN to estimate the channel parameters for different cases. Note that the curve-fitting technique requires simulation data but a trained ANN does not require any simulation data, i.e., required inputs are the system parameters such as \( d, r_{\text{Tx}}, D, \) and \( r_{\text{Rx}}. \)

### IV. Results and Analysis

#### A. Performance Metrics and Parameters

For the performance metrics, we mainly use root mean squared error (RMSE) with respect to simulation data in terms of number of received molecules. First, we give the received signal and signal-to-interference (SIR) plots for the example cases. We then present the average RMSE over different cases.

Common system parameters for simulation, training (TDS), and validation (VDS) datasets are presented in Table I. From the given datasets, each of the VDS and TDS have \( 5 \times 3 \times 3 = 135 \) different cases, making a total of 270 cases. Each simulation case is replicated 500 times to estimate the mean received signal at the receiver side.

What is of prime interest for modeling an MCvD channel is to model, as noted above, the received signal (i.e., the number of hitting molecules until time \( t \)). To model the received signal, we utilize curve-fitting and artificial neural network techniques. In the performance figures, we cannot present all 135 cases, but we offer some example scenarios and average RMSE plots.

#### B. Received Signal Analysis

In Fig. 3 the received signal is plotted for simulation data, curve-fitting, and ANN techniques. Note that the ANN technique requires no simulation data, while the curve-fitting method does. After training an ANN, we estimated channel model parameters for the validation data by giving only the system parameters as input. For the primitive model that is given in (6), a single channel model parameter was estimated. On the other hand, for the enhanced model that is given in (7), three channel model parameters were estimated.

It can be clearly seen that the enhanced model fits the simulation data better than does the primitive model. At the peak and the tail part of the received signal, we observe that the enhanced model outperforms the primitive model and the point transmitter formulation given in (5). The second observation suggests that, with increased distance, the estimation performance of the received signal is improved. This observation is also supported with the RMSE plots in Fig. 5. Another observation is that the trained ANNs generalize the fitted model parameters well. In other words, we observe that the ANN curves are close to the curves that are produced by curve fitting. Without knowing the simulation data (by using only the system parameters as inputs), the trained ANN estimates the channel model parameters well.

#### C. SIR Analysis

For a more in-depth analysis we used the following metric, SIR, formulated as:

\[
SIR(t) = \frac{F_{\text{hit}}^{3D}(t)}{F_{\text{hit}}^{3D}(t_{\text{end}}) - F_{\text{hit}}^{3D}(t)}
\]

where we can also substitute \( F_{\text{hit}}^{3D}(t) \) with the channel model functions given in (6) and (7). This metric basically represents the ratio of the cumulative number of received molecules until time \( t \) and the number of interference molecules. SIR plots are important since they show the performance of modeling the ratio of desired signal and interference. In Fig. 4 SIR is plotted for the simulation data, curve-fitting, and ANN techniques.
The enhanced model estimates the received signal better than the primitive model. Moreover, the longer the distance the better the channel parameter estimation performance. For $d = 9\,\mu m$ case, the enhanced model is very close to the simulation data. However, the primitive model with one scaling factor does not adequately model the simulation data. Again, we see that the trained ANN generalizes the fitting method well for both of the cases (i.e., fitting and ANN curves are overlapping for both models). After $d = 7\,\mu m$, ANN performs close to the simulation data.

### D. RMSE Analysis

To better understand the performance of the estimation technique, we evaluated the mean RMSE of cases with respect to (wrt) simulation data in terms of the number of received molecules until time $t$. For RMSE analysis, we grouped results with respect to distance and $r^{Rx}$ so that each group was the average of nine cases.

In Fig. [5] the RMSEs of the channel parameter estimation methods are presented for ANNs with primitive and enhanced models. The first observation suggests that the estimation performance gets better with increasing distance. Moreover, the RMSE of the enhanced model is significantly lower than that of the primitive model. Note that the ANN technique requires no simulation data, utilizing only the system parameters as input to estimate the channel parameters. The primitive model scales the formulation of point transmitter case and the larger $r^{Rx}$ case deviates more from the point transmitter case, as the probability rises of receiving obstructed and reflected molecules. Therefore, for the primitive model, performance of the estimation of the received signal is better in terms of RMSE for smaller $r^{Rx}$.

### V. Conclusion

In this work, we developed a novel technique to model the received signal in MCvD with a spherical transmitter. In the literature, a point transmitter is assumed for the tractability of the mathematical derivations in a first-hitting process framework. Approaching the problem from a unique perspective, we utilized an artificial neural network technique and a model function for the number of received molecules. After training an ANN, we were able to ask the ANN to estimate the channel model parameters for different system setups. Our proposed technique has promising results for modeling the number of received molecules until time $t$. We observed that the proposed
technique models the received signal and SIR more effectively for longer distances. The proposed technique may be utilized to model an MCVD channel in other studies that assumes a spherical transmitter instead of a point transmitter.

ACKNOWLEDGMENT

This research has been supported by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the “ICT Consilience Creative Program” (IITP-R0346-16-1008) supervised by the IITP (Institute for Information & Communications Technology Promotion) and by the Basic Science Research Program (2014R1A1A1002186) funded by the MSIP, Korea, through the National Research Foundation of Korea.

REFERENCES

[1] N. Farsad, H. B. Yilmaz, A. W. Eckford, C.-B. Chae, and W. Guo, “A comprehensive survey of recent advancements in molecular communication,” IEEE Commun. Surveys Tuts., vol. 18, no. 3, pp. 1887–1919, 2016.

[2] T. Nakano, A. W. Eckford, and T. Haraguchi, Molecular communication. Cambridge University Press, 2013.

[3] I. F. Akyildiz, J. M. Jornet, and M. Pierobon, “Nanonetworks: a new frontier in communications,” Commun. ACM, vol. 54, no. 11, pp. 84–89, Nov. 2011.

[4] W. Guo, T. Asyhari, N. Farsad, H. B. Yilmaz, B. Li, A. Eckford, and C.-B. Chae, “Molecular communications: channel model and physical layer techniques,” IEEE Wireless Commun. Mag., vol. 23, no. 4, pp. 120–127, 2016.

[5] S. Hiyama, Y. Moritani, T. Suda, R. Egashira, A. Enomoto, M. Moore, and T. Nakano, “Molecular communication,” in Proc. NSTI Nanotechnol. Conf. (NanoTech), 2005, pp. 391–394.

[6] L. Parcerisa Gine and I. F. Akyildiz, “Molecular communication options for long range nanonetworks,” Elsevier Comput. Netw., vol. 53, no. 16, pp. 2753–2766, Nov. 2009.

[7] M. Gregori and I. F. Akyildiz, “A new nanonetwork architecture using flagellated bacteria and catalytic nanomotors,” IEEE J. Sel. Areas Commun., vol. 28, no. 4, pp. 612–619, May 2010.

[8] W. C. Agosta, Chemical communication: the language of pheromones. Macmillan, 1992.

[9] H. C. Berg, Random walks in biology. Princeton University Press, 1993.

[10] P. Lio and S. Balasubramaniam, “Opportunistic routing through conjunction in bacteria communication nanonetwork,” Elsevier Nano Commun. Netw., vol. 3, no. 1, pp. 36–45, Mar. 2012.

[11] M. S. Kuran, H. B. Yilmaz, T. Tugcu, and B. Ozerman, “Energy model for communication via diffusion in nanonetworks,” Elsevier Nano Commun. Netw., vol. 1, no. 2, pp. 86–95, June 2010.

[12] K. V. Srinivas, A. W. Eckford, and R. S. Adve, “Molecular communication in fluid media: The additive inverse Gaussian noise channel,” IEEE Trans. Inf. Theory, vol. 58, no. 7, pp. 4678–4692, Jul. 2012.

[13] S. Kadloor, R. S. Adve, and A. W. Eckford, “Molecular communication using Brownian motion with drift,” IEEE Trans. NanoBiosci., vol. 11, no. 2, pp. 89–99, June 2012.

[14] T. Nakano, Y. Okaie, and J.-Q. Liu, “Channel model and capacity analysis of molecular communication with Brownian motion,” IEEE Commun. Lett., vol. 16, no. 6, pp. 797–800, Jun. 2012.

[15] H. B. Yilmaz, A. C. Heren, T. Tugcu, and C.-B. Chae, “Three-dimensional channel characteristics for molecular communications with an absorbing receiver,” IEEE Commun. Lett., vol. 18, no. 6, pp. 929–932, Jun. 2014.

[16] H. B. Yilmaz and C.-B. Chae, “Simulation study of molecular communication systems with an absorbing receiver,” Elsevier Simul. Model. Pract. Theory, vol. 49, pp. 136–150, Dec. 2014.

[17] ——, “Arrival modelling for molecular communication via diffusion,” IET Electron. Lett., vol. 50, no. 23, pp. 1667–1669, Nov. 2014.

[18] A. Akkaya, H. B. Yilmaz, C.-B. Chae, and T. Tugcu, “Effect of receptor density and size on signal reception in molecular communication via diffusion with an absorbing receiver,” IEEE Commun. Lett., vol. 19, no. 2, pp. 155–158, Feb. 2015.