Solving Optimal Neural Layout by Gibbs Sampling

Tao Jiang[*]   **Peng Xu**[§]   **Pamela A. Abshire**[§]   **John S. Baras**[*]

Institute for Systems Research and Department of Electrical and Computer Engineering
University of Maryland, College Park, MD 20742 USA
{tjiang, pxu, pabshire, baras}@isr.umd.edu

Abstract

Neural systems of organisms derive their functionality largely from the numerous and intricate connections between individual components. These connections are costly and have been refined via evolutionary pressure that acts to maximize their functionality while minimizing the associated cost. This tradeoff can be formulated as a constrained optimization problem. In this paper, we use simulated annealing, implemented through Gibbs sampling, to investigate the minimal cost placement of individual components in neural systems. We show that given the constraints and the presumed cost function associated with the neural interconnections, we can find the configuration corresponding to the minimal cost. We restrict the mechanisms considered to those involving incremental improvement through local interactions since real neural systems are likely to be subject to such constraints. By adjusting the cost function and comparing with the actual configuration in neural systems, we can infer the actual cost function associated with the connections used by nature. This provides a powerful tool to biologists for investigating the configurations of neural systems.

Keywords: Neural systems, optimal layout, simulated annealing, Gibbs sampler, Markov random field

1. Introduction

Connections between neurons in neural systems of organisms play a critical role in shaping their functionality. The connections, which we also call wires, require significant resources such as space, power, and development time. Minimizing the wiring cost while achieving the required functionality confers survival advantages to the organism. Recent work [2, 5] has suggested that the actual layout of neural systems might be the result of wiring cost minimization. Through brute-force search using a linear cost function, it has been shown that the actual ordering of the ganglia of *Caenorhabditis elegans* minimizes the total wire length [2]. For a quadratic cost function, an analytic solution exists for the optimal layout problem [3]. The solution gives the same ordering for *C. elegans* ganglia as the actual layout for all the ganglia except one. In a very different neural system, the optimal wiring solution for the prefrontal cortical area in macaque shows similar patterns of spatial arrangement as the actual ones [5].

So far, most prior research on the neural placement problem found the optimum solution using either exhaustive search [2, 5] or analytic optimization techniques [3]. Note that the possible number of alternative layouts explodes as the number of neurons grows. Therefore brute force searches are impossible even for moderately sized systems. On the other hand, analytic optimization techniques can only provide an exact solution for a few types of cost functions, such as the quadratic cost function in [3]. However, it is important to explore the possibility of other cost functions, since we have no *a priori* reason to believe that any specific function is closest to the one used by nature.

Optimization is universal in all biological systems. For instance, swarms of bacteria, insects, and animals yield sophisticated collective behaviors based only on simple local interactions. In neuronal networks, collective behavior during development achieves optimal placement as components interact through local interconnections. Swarm systems generally involve large number of individuals. Thus scalability and computational complexity are crucial in swarms as well as in neural systems. A stochastic approach for large swarm systems has been proposed [1], where the interconnection of nodes are modeled as Markov random fields (MRFs) and the movement of each node is controlled by simulated annealing with Gibbs sampling. This approach can yield the node configuration corresponding to the global potential minimum. Inspired by the emergent behavior of large swarm systems from local interactions, we model the

neuronal network as a MRF. We design Gibbs potential functions corresponding to the wiring cost function and use simulated annealing with Gibbs sampling to find the neural layout with minimal wiring cost. In this way we can compare the configurations resulting from different cost functions with the actual layout. This may provide some insight into the actual cost functions used by nature.

Furthermore, our method more closely parallels the actual optimization process that occurs during the evolution of biological systems, so it may offer insights into this process as well. Although previous work indicates that neural systems might be optimized to minimize wiring cost, there is limited understanding about how this optimization is actually implemented in such complicated systems with huge numbers of components. It is hard to imagine that nature arrived at this solution by brute-force search or by employing a simple quadratic cost. It is more reasonable to assume that nature uses gradual adjustments and that optimality emerges through millions of years of evolution. Our proposed optimization method using incremental updates based on local interactions is one example of such a possible scheme (or process).

This paper is organized as follows: in Section 2, the neural wiring minimization problem is formulated; in Section 3, Markov random field, Gibbs sampling and simulated annealing are introduced; in Section 4, the algorithm and simulation results on neural wiring cost minimization are described; in Section 5, we summarize our contributions.

2. Optimal neural placement problem and cost function

The functionality of a neural system arises in large part from its connections. In the optimal neural layout problem, we assume that the connections necessary for a specific functionality are known, including both internal connections (connections among the neurons) and external connections (connections to neurons external to the neural network). The task is to find the placement of the neurons that minimize the cost associated with the connections.

We abstract the neuronal network as a non-directed weighted graph. Nodes of the graph represent individual components of the neuronal network. Depending on the level of the network, they can be neurons, clusters of neurons such as ganglions, or sub-neuronal networks. For simplicity, we use these interpretations interchangeably. Edges of the graph represent connections; the weight of each edge represents the connection strength. In addition, there are edges from the nodes of the network to external nodes which correspond to the external components connected with the neuronal network. The positions of external nodes are fixed.

The graph is specified by the adjacency matrix, A, where element A_{rs} gives the connection strength between nodes r and s. The connection to external nodes is specified by the matrix B, where element B_{rt} gives the connection strength between node r and external node t. The total cost of the wires is given by

$$C = \frac{1}{2}\sum_{r,s} A_{rs} f(x_r, x_s) + \sum_{r,t} B_{rt} g(x_r, y_t)$$

where x_r, x_s are the positions of the nodes, y_t is the position of the external node, $f(x_r, x_s)$ is the wiring cost between nodes r and s, and $g(x_r, y_t)$ is the wiring cost between node r and external node t. To solve the optimal layout problem, we search for the node positions giving minimum wiring cost.

The cost associated with each wire could arise from its volume, metabolic requirements, signal transmission, or development guidance. The farther apart two connected neurons are, the higher the wiring cost. However, it is not clear what exact function of the distance the cost should be. Brute force enumeration has been used to find the optimal layout for a linear cost function. It has also been argued that the cost should depend on $L^{3n/(n+2)}$ due to the trade-off between wire volume and signal propagation delay, where L is the length of the wire and n is a positive number [3]. For n equal to or greater than 1, the cost function lies between linear and cubic functions of the length. So the cost function can be written as

$$C = \frac{1}{2}\sum_{r,s} A_{rs} |x_r - x_s|^{\gamma} + \sum_{r,t} B_{rt} |x_r - y_t|^{\gamma}$$

with $\gamma \in [1,3]$. In [3], the quadratic cost function ($\gamma = 2$) was chosen because it provides an analytic solution to the wiring cost minimization problem. For other cost functions, there are no analytic solutions, and solving the optimal layout problem is complicated due to the large number of possible spatial arrangements of all the components.

Using our proposed method for solving the optimal layout problem, we are able to investigate many different cost functions. By comparing the resulting solutions with the actual layout we hope to find tight estimates of the actual cost functions used by nature.

3. Local interaction and global optimum

As mentioned in previous work [2], the nervous system is far from a completely interconnected network, where "everything is connected to everything". In the connectivity matrix of *C. elegans*

ganglia, about half of the matrix elements remain empty, i.e., about half of the ganglia pairs do not interconnect. In addition, connections tend to link pairs of ganglia which are adjacent or at least nearby. Information is exchanged only among the interconnected neurons, so each neuron obtains information from a subset of the entire system and usually from the nearby portion. This leads naturally to the question: how does a real physical system reach an optimum using only partial or local information exchanges?

A disadvantage for optimization methods that depend only on local interactions is that they are very easily trapped in local optima, and they are likely to misinterpret local optima as global ones. In order to avoid these local optima, we adopt a stochastic approach which is based on the theory of Markov random fields (MRFs) and simulated annealing with Gibbs sampling and can yield the desired configurations corresponding to global minima of the cost function using only local interactions [1]. We first introduce MRFs and the Gibbs distribution.

3.1 MRFs and Gibbs distribution

Define the neighborhood of a site s as $\mathcal{N}_s = \{r | r$ and s *interconnect* $r \neq s \}$. In MRFs, the value corresponding to s is independent of non-neighbors given the values of all its neighbors. In the neural placement problem, the *sites* are the neurons and the corresponding random variables are their positions (to be determined), which are denoted as X_s for neuron s, where $s \in S$. Then the conditional probability of the MRF can be represented as: $\forall s \in S$

$$P(X_s \mid X_r, r \neq s) = P(X_s \mid X_r, r \in \mathcal{N}_s)$$

i.e., the conditional probabilities depend only on neighbors.

For local interactions, the well-known Hammersley-Clifford theorem proves the equivalence between a MRF and the Gibbs distribution, whose joint probability distribution is of the form

$$P(X = x) = e^{-\frac{U(x)}{T}} \Big/ Z$$

where T is the temperature (discussed further in Sect. 3.2), $U(x)$ is the potential, and Z is the normalization constant. In the neural placement problem, the potential $U(x)$ is naturally set to be the cost function (for simplicity, only internal costs are considered, but it is easy to extend to include external costs):

$$U(x) = \frac{1}{2} \sum_{r,s} A_{rs} \mid x_r - x_s \mid^{\gamma} \qquad (1)$$

Then if the positions of any neurons $r \neq s$ are fixed as x_r, the probability that the position of s is z is defined as

$$P(X_s = z \mid X_r = x_r, r \in S, r \neq s)$$

$$= \frac{\exp\left(-U(x_1, \ldots, x_{s-1}, z, x_{s+1}, \ldots, x_{|S|})/T\right)}{\sum_{z'} \exp\left(-U(x_1, \ldots, x_{s-1}, z', x_{s+1}, \ldots, x_{|S|})/T\right)}$$

$$= \frac{\exp\left(-\dfrac{1}{T} \sum_{r \in \mathcal{N}_s} A_{sr} \mid z - s_r \mid^{\gamma}\right)}{\sum_{z'} \exp\left(-\dfrac{1}{T} \sum_{r \in \mathcal{N}_s} A_{sr} \mid z' - s_r \mid^{\gamma}\right)} \qquad (2)$$

where the last equality is obtained by substituting Eqn. (1). Notice that the conditional probability only depends on the neighbors of s. This also verifies the equivalence of MRFs and Gibbs distributions. Next, we introduce the stochastic method that achieves the desired minimum cost based on the MRF property.

3.2 Simulated annealing with Gibbs sampling

We allow local changes of the neuron positions obtained by randomly sampling the conditional distribution of Eqn. (2), where the local conditional distributions are dependent on a global control parameter T called "temperature". At low temperatures the distributions concentrate on states that decrease the cost function, whereas at high temperatures the distribution is essentially uniform. T is initially large, so the process avoids becoming trapped in local minima. Then temperature is gradually lowered and neural positions are iteratively adjusted to minimize the cost function. This gradual reduction of temperature simulates "annealing" and has been shown [4] to converge to the global maxima of the Gibbs distribution, which corresponds to the placements with minimum cost. The whole stochastic process works as follows:

1. *Initialization*: Pick a cooling schedule for T and randomly select the initial position of each neuron.
2. *Annealing*: At each temperature, visit all the neurons a certain number of times. Update the position of each neuron in turn. When visiting s, fix the positions of all other neurons $r \neq s$, and change the position of s to z with the probability defined in Eqn. (2).
3. *End*: Repeat the 2nd step until the cooling schedule ends.

4. Algorithm implementation and results

We implemented our algorithm in Matlab using the interconnection matrix for *C. elegans* ganglia provided in [2]. We assume that the neurons and all
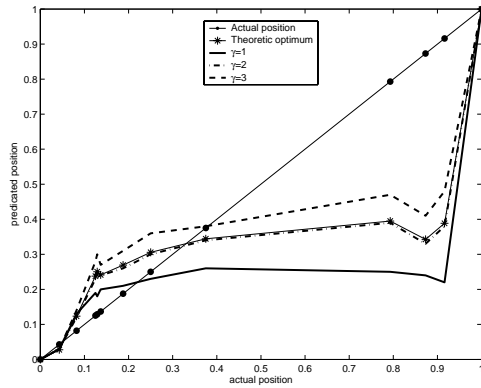
Fig. 1: Optimal placement under different cost functions ($\gamma = 1,2,3$). Dots on the diagonal line are the actual positions of 11 C. *elegans* ganglia with normalized length. The line with stars is the exact solution with quadratic cost from [3].The thick lines are the results of Gibbs sampling.

Fig. 2: Optimal placement with different external and internal cost exponents. ($\gamma_{inter} = 1, \gamma_{exter} = 2$)

external sensors and organs are located in a unit length line because of the worm's large aspect ratio. The head and tail are at positions 0 and 1 respectively. For the software implementation, we divided the unit line into 100 small intervals of equal length. Each interval represents a position, and intervals are ordered from 1 to 100. In order to emphasize the **locality** of our algorithm, the candidate positions for a neuron to move at each temperature are only those at most 2 positions away, i.e., if a neuron is at z, then it can only choose positions from the set $\{z-2, z-1, z, z+1, z+2\}$ for the next iteration.

We compare the exact solution and our result for the quadratic cost function ($\gamma = 2$) in Fig 1. Our result approaches the exact solution provided in [3]. In Fig. 1, results with different cost exponents are also shown. We observe that in all three cases, the order of the neurons is nearly the same as the actual order except for one or two ganglia that are slightly different. This verifies that wire length is an important factor in neuron placement. Moreover, the solution for a linear cost function performs slightly better, especially for neurons located near the head. However, for ganglia near the tail, none of the three cases gives good results. Tail side ganglia are all shifted toward the head.

By examining the connectivity matrix, we observed that tail side neurons have relatively strong connections with external sensors or organs that are located on the tail. So we modified our cost function to include an external cost with high penalty. Figure 2 shows the solution which results from using the cost function

$$c(x) = \frac{1}{2}\sum_{r,s} A_{rs} \mid x_r - x_s \mid^{\gamma_{inter}} + \sum_{r,t} B_{rt} \mid x_r - y_t \mid^{\gamma_{exter}}$$

where $\gamma_{inter} = 1$ and $\gamma_{exter} = 2$. This cost function gives a much better estimate of the actual positions, especially for ganglia near the tail.
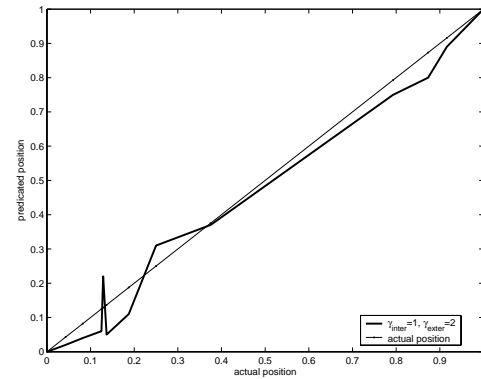
5. Conclusions

We introduce a stochastic method based on simulated annealing with Gibbs sampling for the neural layout optimization problem, which is computationally feasible and can handle all kinds of cost functions. It also provides a new way to explore the detailed evolution of biological systems. Using this method, we defined a new cost function that distinguishes between internal and external connections. This cost function estimates actual neural positions much better than previous methods.

We are also considering two-dimensional models of neural interconnection, such as the prefrontal cortical area in macaques or cats. In future work we will investigate more general cost functions that achieve better predictions of neural placement.

6. References

[1] J. S. Baras, and X. Tan, "Control of Autonomous Swarms Using Gibbs Sampling," *Proc. Of 43rd IEEE Conference on Decision and Control*, pp. 4752-4757, 2004.

[2] C. Cherniak, "Component Placement Optimization in the Brain," *The Journal of Neuroscience*, 14(4):2418-2427, April 1994.

[3] D. B. Chklovskii, "Exact Solution for the Optimal Neuronal Layout Problem," *Neural Computation*, 16, pp. 2067-2078, 2004.

[4] S. Geman, and D. Geman, "Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images," *IEEE Tran. on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721-741, 1984.

[5] V. A. Klyachko, and C. F. Stevens, "Connectivity Optimization and the Positioning of Cortical Areas," *Proc. Of the National Academy of Sciences*, 100(13):7937-7941, 2003.