ISSN:0975-9646

Deriving Serial and Parallel Spikes Data in Neural Science Using Data Mining Techniques

Srinivas Bobba[1], Jammi Ashok[2]
[1]*Dept. of IT,Geethanjali college of Engg. & Tech,Hyderabad*
[2]*Dept. of CSE, Gurunanak Institute of Technology,Hyderabad*

Abstract – In this paper we introduce the notion of mining for frequent episodes under certain temporal constraints; the structure of these temporal constraints is motivated by the application. We present algorithms for discovering serial and parallel episodes under these temporal constraints.

Keywords: - Spike data Analysis, Temporal data mining, frequent episodes, and temporal constraints

1. INTRODUCTION

Over the last couple of decades, biology has thrown up many interesting and challenging computational problems. For example, the problem of understanding genome data and protein function has motivated development of many computational and statistical techniques leading to the creation of the interdisciplinary area of Bioinformatics. One of the main driving forces in this case is the availability of large amounts of data, from gene or protein sequencing experiments, and the consequent need for efficient techniques to analyze the data to arrive at reasonable and useful inferences. To solve these computational problems, some techniques developed in other contexts e.g., Hidden Markov Models, Dynamic Programming have proved to be quite suitable, after some modifications. In this project, we focus on an equally challenging computational problem in another sub area of biology, namely neuroscience. We look at the problem of analyzing multi-neuronal spike train data and suggest that certain techniques from the field of Temporal Data mining.

2. LITERATURE SERVEY

Over the past twenty years or so, increasingly better methods are becoming available for simultaneously recording the activities of many neurons. By using techniques such as micro electrode arrays, imaging of currents, voltages, and ionic concentrations etc., spike data can be recorded simultaneously from hundreds of neurons. Vast amounts of such data are now routinely gathered from different neuronal systems. Each recording session contains data with tens of thousands of spikes. Such multi-neuronal spike train data can now be obtained from neuronal cultures from brain slices. Such spike train data is a mixture of the spiking activities of individual neurons as well as correlated spiking activity due to interactions or connections among neurons.

It is well accepted that the brain's computation relies on spatiotemporal activity of neural networks. In particular, there is growing evidence of the importance of continuously and precisely timed spiking activity. Therefore, it is important to characterize memory states in terms of spike-timing patterns that give both reliable memory of firing activities and precise memory of firing timings. The relationship between memory states and spike-timing patterns has been studied empirically with large-scale recording of neuron population in recent years. Here, by using a recurrent neural network model with dynamics at two time scales, we construct a dynamical memory network model which embeds both fast neural and synaptic variation and slow learning dynamics. [1]

A state vector is proposed to describe memory states in terms of spike-timing patterns of neural population, and a distance measure of state vector is defined to study several important phenomena of memory dynamics: partial memory recall, learning efficiency, learning with correlated stimuli. We show that the distance measure can capture the timing difference of memory states. In addition, we examine the influence of network topology on learning ability, and show that local connections can increase the network's ability to embed more memory states. Together these results suggest that the proposed system based on spike-timing patterns gives a productive model for the study of detailed learning and memory dynamics. [1]

Recurrent neural networks of the brain compute information through complexly spatiotemporal neural activity. Recent experimental observations and theoretical studies have proposed that spike-timing patterns (STPs) in the range of a few hundred milliseconds play a fundamental role in sensory, motor and high-level cognitive behaviors such as learning and memory. For instance, songbirds, one of the most studied neural systems, learn and memorize the crystallized song composed by precise individual syllables as STPs. Traditionally, firing rate is used to describe the activity of single neurons and neural networks. However, a memorized song as a STP contains not only firing activity, i.e., whether neurons fire or not, but also firing timings, i.e., when neurons fire. Therefore, memory has to be both reliable in firing and precise in timing. Firing rate as a average measure is reliable but not precise. Thus the question is how to capture the precise timing of memory from STPs.

Recent experimental data of hundreds of spike trains from multi-electrode recording have identified repeated or periodic STPs. There is a great interest in such a STP code in neural circuits. Neurons in vitro produce a STP in response to an external stimulus. However, neurons in vivo are modulated by local oscillatory neural activities and top-down inputs. In a cortical circuit, precise STPs thus reflect the interaction between internally generated activity and sensory information. On the other hand, memory states are global dynamical behaviors of the cortical network emerged from relatively simple neural and synaptic dynamics. It has shown that several different dynamical regions for the spontaneous network activity generated by Poisson background inputs can be identified. However, dynamical behavior of neural networks in response to external stimuli is less well studied due to the difficulty of the mathematical description of nonlinear high dimension dynamical system. Thus the essential question is how to construct a global description of network states in terms of STPs, which is less dependent of the existence of background spiking noise and external inputs. [1]

In this work, we address these questions by simulating a two time-scale biologically realistic neural network with dynamics evolving at two time-scales: the fast scale of neurons and synapses and slow scale of homeostatic presynaptic-dependent synaptic scaling. After training, the network converges to a stable state with a spare neural trajectory as a STP. By proposing a state vector for the STP induced by each stimulus, we show the distance of state vectors can be used to characterize learning process and several important phenomena of memory dynamics: partial memory recall, learning efficiency, learning with correlated stimuli. Specifically, we examine the influence of network topology on leaning ability, and show that local connections can increase the network's ability to embed more memory states.

We also show that distance measure can capture the timing difference of memory states formed in partial memory recall tasks and correlated-stimuli learning tasks. However, firing rate and correlation coefficient fail to differentiate these similar memories. Together these results suggest that the proposed system based on spike-timing patterns gives a productive model for the study of detailed learning and memory dynamics.[1]

Understanding the functioning of a neural system in terms of its underlying circuitry is an important problem in neuroscience. Recent developments in electrophysiology and imaging allow one to simultaneously record activities of hundreds of neurons. Inferring the underlying neuronal connectivity patterns from such multi-neuronal spike train data streams is a challenging statistical and computational problem. This task involves finding significant temporal patterns from vast amounts of symbolic time series data. In this paper we show that the frequent episode mining methods from the field of temporal data mining can be very useful in this context. In the frequent episode discovery framework, the data is viewed as a sequence of events, each of which is characterized by an event type and its time of occurrence and episodes are certain types of temporal patterns in such data.

Here we show that, using the set of discovered frequent episodes from multi-neuronal data, one can infer different types of connectivity patterns in the neural system that generated it. For this purpose, we introduce the notion of mining for frequent episodes under certain temporal constraints; the structure of these temporal constraints is motivated by the application. We present algorithms for discovering serial and parallel episodes under these temporal constraints. Through extensive simulation studies we demonstrate that these methods are useful for unearthing patterns of neuronal network connectivity.[1]

Over the last couple of decades, biology has thrown up many interesting and challenging computational problems. For example, the problem of understanding genome data and protein function has motivated development of many computational and statistical techniques leading to the creation of the interdisciplinary area of Bioinformatics. One of the main driving forces in this case is the availability of large amounts of data, from gene or protein sequencing experiments, and the consequent need for efficient techniques to analyze the data to arrive at reasonable and useful inferences. To solve these computational problems, some techniques developed in other contexts (e.g., Hidden Markov Models, Dynamic Programming) have proved to be quite suitable, after some modifications. In this paper, we focus on an equally challenging computational problem in another sub area of biology, namely neuroscience.

We look at the problem of analyzing multi-neuronal spike train data and suggest that certain techniques from the field of Temporal Data mining are attractive here. Neurons form the basic computing elements of brain and hence, gaining an understanding of the coordinated behavior of groups of neurons (at different levels of organization) is essential for gaining a principled understanding of brain function. Thus, one of the important problems in neuroscience is that of understanding the functioning of a neural tissue in terms of interactions among its neurons. Many neurons communicate with each other through characteristic electric pulses called action potentials or spikes. Hence one can study the activity of a specific neural tissue by gathering data in the form of sequences of action potentials or spikes generated by each of a group of potentially interconnected neurons. Such data is known as multi-neuronal spike train data.[2]

The brain or the nervous system consists essentially of a vast network of neurons. The neuron may be regarded as the basic computing element in the nervous system. Each neuron is connected to many others through what are known as synapses. Synapses, through electric and chemical means, allow neurons to signal to each other in the sense that the output of one neuron can become input to another through the synapse that connects them. In a good majority of all neurons, the output of a neuron is in the form of what is called an action potential. An action potential is an electrical signal of a short duration (typically less than 1 ms) with a characteristic shape. For most purposes of analysis, this can be regarded as a short pulse and hence is also referred to as a spike. After generating an action potential, a neuron can not immediately generate another spike because it needs some regenerative time. This time period is called refractory period and in many cases it is in the range of 1 milli second (ms). Over short durations of time the spiking activity of a neuron can be well modeled by a Poisson process whose rate depends on the current state of the neuron. [1]

The spikes output by one neuron reach the input terminals of other neurons through the synapses that interconnect them. Each neuron has many such synapses and based on the amount of input it receives like this, it may then fire an action potential or a spike. The functioning of the nervous system is essentially due to this coordinated activity of many neurons. The system is stochastic and neurons would also be spiking randomly. The signal transmission through synapses takes some time and thus there are characteristic delays associated with each synapse. Also different synapses may have different efficacies in effecting spikes from the receiving neurons.

Experimental studies for understanding brain function span a wide range of organizational levels. At one end are studies aimed at understanding the functioning of single neurons through electro physiological recordings while at the other end, using techniques such as fMRI, one studies interactions among large brain regions. Our interest in this paper is in experimental techniques at an intermediate level of organization where one is interested in understanding how groups of a few hundred neurons act in a coordinated manner to generate specific functions. For this, as mentioned earlier, one obtains simultaneous recordings of the spikes generated by a group of interacting neurons. By simultaneous recording we mean that the times of spikes of all neurons are referenced with respect to a common time origin and hence the data is suitable for studying temporal interactions among the neurons. [3]

There has been a lot of work on experiments for simultaneously recording the activities of hundreds of neurons for gaining a better understanding of the functional interactions among neurons in a neural tissue. The recording techniques fall into three broad categories. In the first category are recordings from cultured neurons or brain slices using Micro electrode arrays (MEAs). A typical MEA setup for this consists of 8 × 8 grid of 64 electrodes with inter-electrode spacing of about 100 microns. This allows stimulation of the neural tissue and recording of the resulting spikes using the same set of electrodes. In the second category are recordings from intact animals using MEAs and other probes?

In the third category are imaging techniques using voltage sensitive dyes and indicators for ions such as Ca++. One of the most exciting recent developments is the incorporation of ion-selective pores into neurons of behaving animals. This allows simultaneous stimulation and recording with milli-second precision using light at various wavelengths. All these technologies now allow for gathering of vast amounts of data, using which one wishes to study connectivity patterns and microcircuits in neural systems. [2]

In this paper our interest is in techniques for analyzing the data that is in the form of spike trains. To obtain such spike trains from the recorded data, one needs a lot of signal processing and data preprocessing techniques. For example, in MEA experiments, the raw data is in the form of voltage or current signals from each of the electrodes, recorded at a suitably high sampling rate. By employing appropriate signal processing techniques one has to first reliably locate all spiking events. Even after this, what we have are spike events in each channel or each electrode.

Since the micro electrode array is regular while the neuronal tissue is not, each electrode may be picking up signals from many neurons with different efficacies. If we want the final data as spike events generated by individual neurons then we have to do what is called spike sorting. Many techniques have been suggested for processing the raw signals to obtain spikes and to do spike sorting and there is need for better algorithms for these problems. Here we will not review any of these techniques because our interest is in analysis methods that look at spike trains to infer connectivity patterns. [4]

The field of multi-neuronal data analysis has a long history. A major goal of such neural data analysis is to characterize how neurons that are part of an ensemble interact with each other. Statistical analysis of spike train data was pioneered and the recent review addresses all the statistical issues in this area. The review summarizes three decades of methodology development in this area and eloquently lays out future challenges. Many of the current methods essentially use information obtained from cross correlation among spike trains that are shifted in time with respect to one another. For example, one can compute what is called a joint peri-stimulus time histogram (JPSTH) which is a two dimensional histogram that displays the joint spike count of two neurons at different time lags (for a specific binning on the time axis). [4]

There are other methods based on analyzing time-shifted spike trains for detecting repeated patterns of firing of a few neurons with constant time lags. Given some specific patterns there are methods to look for matches in the spike train data and assess the statistical significance. For assessing the statistical significance of the detected patterns, one generally employs a null hypothesis that the spike trains are iid Bernoulli processes or uses some resampling methods of generating surrogate data streams to assess significance empirically. Most of these techniques are not efficient for detecting patterns involving more than four or five neurons. Another approach is to employ dimensionality reduction techniques such as PCA and study the data in some appropriate low dimensional feature space.

3. PROPOSED SYSTEM

System Algorithm

Mining Frequent Episodes

1: Generate an initial set of (1-node) candidate episodes (N=1)

2: repeat

3: Count the number of occurrences of the set of (N-node) candidate episodes in one pass of the data sequence

4: Retain only those episodes whose count is greater than the frequency threshold and declare them to be frequent episodes

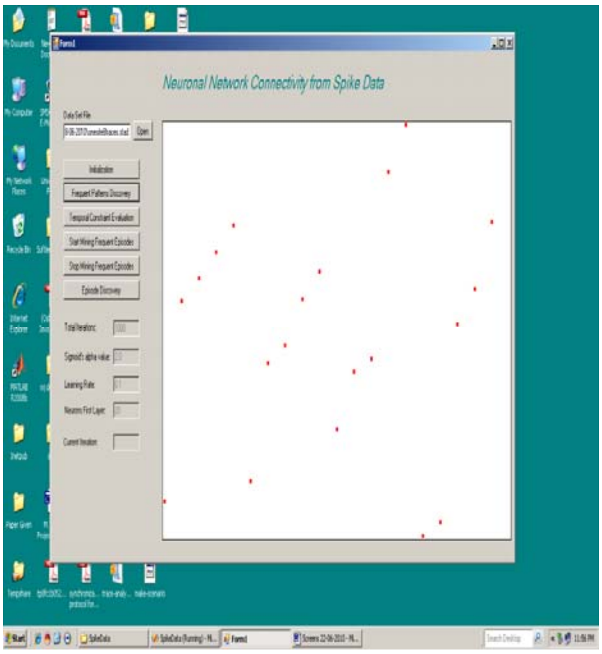
5: Using the set of (N-node) frequent episodes, generate the next set of (N+1-node) candidate episodes

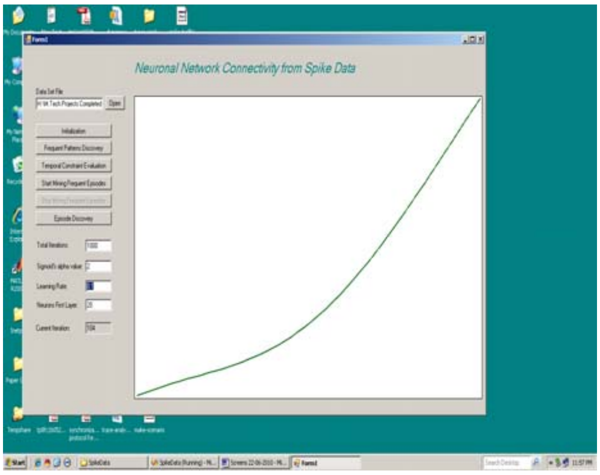
6: until There are no candidate episodes remaining

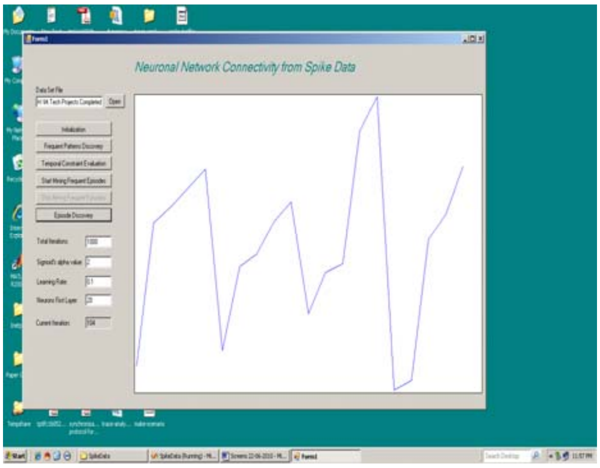
7: Output all the frequent episodes discovered

4. RESULTS

The concept of this paper is implemented and different results are shown below







5. CONCLUSION

Analyzing multi-neuron spike data is a challenging problem of much current interest in neuroscience. With the recent advances in experimental techniques, we can now easily obtain data representing the simultaneous activity of hundreds of neurons. Hence algorithms that can discover significant patterns of co-ordinated spiking activity among neurons would be very useful in making sense of these vast amounts of data. Discovering such patterns would help in understanding the underlying connectivity structure in the neural tissue and to relate it to the function of the nervous system. Such an understanding of the behavior of interacting neurons is very useful in elucidating issues such as learning and memory as well as for applications such as brain computer interfaces.

6. REFERENCES

[1] Bettencourt LM, Stephens GJ, Ham MI, Gross GW: The functional structure of cortical neuronal networks grown in vitro. Phys Rev E 2007, 75:021915.
[2] Takahashi DY, Baccalá LA, Sameshima K: Connectivity inference between neural structures via partial directed coherence. Journal of Applied Statistics 2007, 34:1259-1273.
[3] Makarov VA, Panetsos F, De Feo O: A method for determining neural connectivity and inferring the underlying network dynamics using extracellular spike recordings. J Neurosci Methods 2005, 144:265-279.