On the Time Series K-Nearest Neighbor Classification of Abnormal Brain Activity

Article *in* IEEE Transactions on Systems Man and Cybernetics - Part A Systems and Humans · December 2007

DOI: 10.1109/TSMCA.2007.897589 · Source: IEEE Xplore

3 authors:

Wanpracha Chaovalitwongse
University of Washington Seattle
158 PUBLICATIONS **2,022** CITATIONS
SEE PROFILE

Ya Ju Fan
Lawrence Livermore National Laboratory
17 PUBLICATIONS **182** CITATIONS
SEE PROFILE

Rajesh C Sachdeo
Jersey Shore University Medical Center
76 PUBLICATIONS **4,432** CITATIONS
SEE PROFILE

On the Time Series K-Nearest Neighbor Classification of Abnormal Brain Activity

Wanpracha Art Chaovalitwongse, *Member, IEEE*, Ya-Ju Fan, *Student Member, IEEE*, and Rajesh C. Sachdeo

***Abstract*—Epilepsy is one of the most common brain disorders, but the dynamical transitions to neurological dysfunctions of epilepsy are not well understood in current neuroscience research. Uncontrolled epilepsy poses a significant burden to society due to associated healthcare cost to treat and control the unpredictable and spontaneous occurrence of seizures. The objective of this study is to develop and present a novel classification technique that is used to classify normal and abnormal (epileptic) brain activities through quantitative analyses of electroencephalogram (EEG) recordings. Such technique is based on the integration of sophisticated approaches from data mining and signal processing research (i.e., chaos theory, k-nearest neighbor, and statistical time series analysis). The proposed technique can correctly classify normal and abnormal EEGs with a sensitivity of 81.29% and a specificity of 72.86%, on average, across ten patients. Experimental results suggest that the proposed technique can be used to develop abnormal brain activity classification for detecting seizure precursors. Success of this study demonstrates that the proposed technique can excavate hidden patterns/relationships in EEGs and give greater understanding of brain functions from a system perspective, which will advance current diagnosis and treatment of epilepsy.**

***Index Terms*—Classification, data mining, dynamic time warping (DTW), electroencephalogram (EEG), epilepsy, nearest neighbor.**

I. Introduction

THE HUMAN brain is among the most complex systems that are known to mankind. Neuroscientists seek to understand brain functions through detailed analysis of neuronal excitability and synaptic transmission. However, the dynamical transitions to neurological dysfunctions of brain disorders are not well understood in current neuroscience research. Epilepsy is the second most common brain disorder after stroke yet the most devastating one. The most disabling aspect of epilepsy is the chronic condition of diverse etiologies with the common symptom of spontaneous recurrent seizures, which can be characterized by a chronic medical condition that is produced by temporary changes in the electrical function of the brain. These electrical changes can be captured by electroencephalograms (EEGs), which is a common tool for evaluating the physiological state of the brain. While EEGs offer excellent spatial and temporal resolution to characterize rapidly changing electrical activity of brain activation, neuroscientists understand very little about the seizure development process from EEG data. The unpredictable occurrence of seizures has presented special difficulties regarding the ability to investigate the factors by which the initiation of seizures occurs in humans. If seizures could be predicted, it will revolutionize neuroscience research and provide greater understanding of abnormal intermittent changes of neuronal cell networks that are driven by the seizure development.

Recent advances in data mining and signal processing research for excavating hidden patterns or relationships in massive data (such as EEGs) offer a possibility to better understand brain functions (as well as other complex systems) from a system perspective, which will generally be very useful in medical diagnosis. Thus, the vital step to advance research in seizure prediction is to develop novel techniques that are capable of recognizing and capturing epileptic activity in EEGs before a seizure occurs. As an answer to this question, the discriminant ability to differentiate and classify the preseizure (abnormal) EEG signal is logically a prerequisite of the seizure prediction/warning process. Thus far, none of the current epilepsy studies in the literature is undertaken to develop quantitative classification techniques that can be used to differentiate normal and preseizure EEG signals. The goal of this study is to develop and present a novel EEG classification technique as a theoretical foundation and host of methodologies to enhance the ability to differentiate normal and abnormal EEG signals. Such technique is based on the integration of the brain dynamics, k-nearest neighbor (KNN), and statistical time series analysis.

A. Background: Epilepsy

At least two million Americans and another 40 million people worldwide (1% of population) currently suffer from epilepsy, which is the second most common brain disorder after stroke [1], [2]. In about half of all cases, no cause can be found. In the other half, head injuries, brain tumors, lead poisoning, problems in brain development before birth, and certain genetic and infectious illnesses are the cause of epilepsy. The mainstay of contemporary treatment for epilepsy is pharmacological, which involves chronic administration of drugs, attempting to keep the concentration of those drugs within a therapeutic window. Antiepileptic drugs are commonly used to reduce

Manuscript received April 7, 2006; revised January 27, 2007. This work was supported in part by the National Science Foundation under CAREER Grant CCF 0546574 and in part by Rutgers Research Council under Grant 202018. This paper was recommended by Associate Editor J. Lee.

W. A. Chaovalitwongse and Y.-J. Fan are with the Department of Industrial and Systems Engineering, Rutgers University, Piscataway, NJ 08854 USA (e-mail: wchaoval@rci.rutgers.edu; yjfan@eden.rutgers.edu).

R. C. Sachdeo is with the Department of Pediatrics, Jersey Shore University Medical Center, Neptune, NJ 07754 USA (e-mail: drsachdeo@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMCA.2007.897589

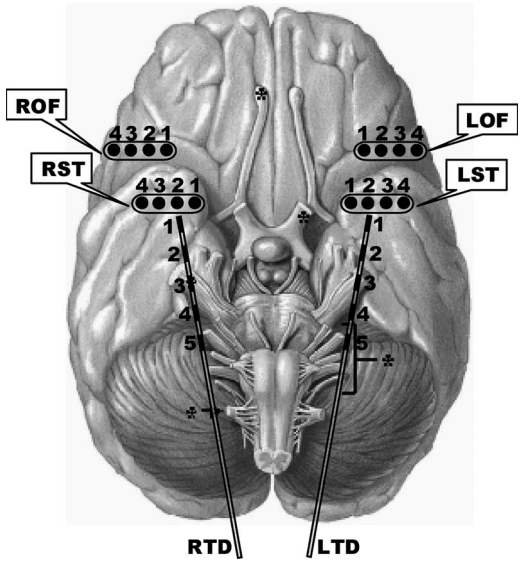

Fig. 1. Inferior transverse views of the brain, illustrating approximate depth and subdural electrode placement for EEG recordings, are depicted. Subdural electrode strips are placed over the left orbitofrontal (LOF), right orbitofrontal (ROF), left subtemporal (LST), and right subtemporal (RST) cortex. Depth electrodes are placed in the left temporal depth (LTD) and right temporal depth (RTD) to record hippocampal activity.

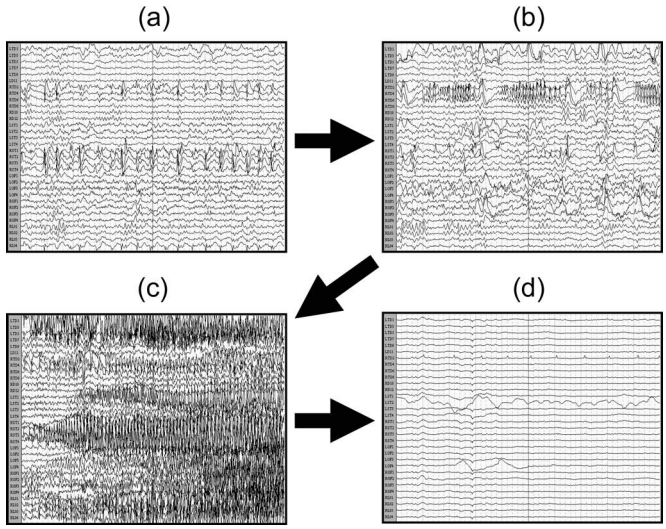

Fig. 2. The 20-s EEG recordings of (a) normal activity, (b) preseizure activity, (c) seizure onset activity, and (d) postseizure activity from patient 1 that were obtained from 32 electrodes. Each horizontal trace represents the voltage that is recorded from the electrode sites that are listed in the left column (see Fig. 1 for anatomical location of electrodes).

the frequency of seizure episodes. However, the limitation of this approach is that drugs have a great deal of side effects, and 25%–30% of epileptic patients remain unresponsive to antiepileptic drug treatment. Alternative treatment is surgical approaches, which attempt to remove the epileptic region of the brain. Nevertheless, surgery is not always feasible and involves the risk of craniotomy. In addition, only 60% of surgical cases are successful, the mean length of hospital stay for patients with intractable epilepsy admitted for invasive EEG monitoring for presurgical candidates ranged from 4.7 to 5.8 days, and the total aggregate costs exceeded $200 million each year [3]. The cost per patient ranged from $4,272 for persons with remission after initial diagnosis and treatment to $138,602 for persons with intractable and frequent seizures [4]. Uncontrolled epilepsy poses a significant burden to society due to associated healthcare cost. The diagnosis and treatment of epilepsy is complicated by the disabling aspect that seizures occur spontaneously and unpredictably due to the nature of the chaotic disorder. The brain activities (both normal and abnormal) are postulated to result from complex systems that follow dynamical transitions and of which the statistical properties depend on both time and space (electrode location). It has been shown that the development of the epileptic state can be considered as changes in the network circuitry of neurons in the brain that produce changes in voltage potential, which can be captured by an EEG. These changes are reflected by wriggling lines along the time axis in a typical EEG recording. For this reason, EEGs have been the main tool for neurologists and neuroscientists that are used to study the epileptogenetic processes and other neurological disorders. A typical electrode montage for the intracranial EEG recordings that are used in our study is shown in Fig. 1. In the emerging view in recent epilepsy research, it is believed that there are four stages that evolved in the seizure process: normal, preseizure, seizure onset, and postseizure [5]–[7]. The 20-s EEG profiles during these stages are illustrated in Fig. 2(a)–(d), respectively.

B. Epilepsy Research: Seizure Prediction

During the past decade, there has been an explosion of interests in applying quantitative methods to seizure prediction research including a measure of the brain dynamics [8], [9], a measure of the relative complexity [10]–[12], a measure of signal similarity in the correlation dimension [13], [14], increasing incidence of energy bursts [15], [16], and multivariate linear discrimination [17]. The studies that are reported in [9], [18]–[22] suggest that seizures are deterministic rather than random, and it may be possible to predict the onset of epileptic seizures based on quantitative analysis of the brain electrical activity through EEGs. Subsequently, the seizure predictability has also been confirmed by several other groups [10]–[17], [23]. The results of these studies indicate that it may be possible to detect state transitions to epileptic seizures based on analysis of the brain electrical activity through EEG signals. It is also believed that a seizure is essentially a reflecting transition of progressive changes of hidden dynamical patterns in EEG. Such a transition that precedes seizures for periods on the order of minutes to hours is detectable in the EEG by the convergence in value of chaos measures among critical electrode sites on the neocortex and hippocampus [21]. This transition has been shown to be detectable through quantitative analysis of the brain dynamics [9], [20]–[24]. The results from our previous studies demonstrated that the spatiotemporal dynamical properties of EEGs manifest patterns corresponding to specific clinical states [9], [20], [25]–[27]. These detectable patterns represent an episode of seizure precursors. These seizure precursors are reflected from the convergence of the brain dynamics from a group of electrode sites during the hour preceding seizures [18], [19], [28], [29].

C. Motivation

Research in seizure prediction is still far from complete in spite of promising signs of the seizure's predictability. While the previous studies demonstrate that preseizure transitions are detectable, the existence of preseizure transitions remains to be further investigated with respect to its specificity and accuracy, i.e., if it only reflects epileptic activity or it also occurs with other brain activity. In addition, the development of a model for the mechanism of generation of epileptic seizures remains a difficult task, and the mechanisms of epileptogenesis are not well understood. Essentially, there is a need for an answer to a fundamental, but extremely crucial, question of whether the brain's normal and preseizure epileptic activities are distinctive or differentiable. Thus far, this question still remains unanswered. Specifically, one needs to demonstrate that normal EEGs differ from preseizure EEGs. In order for one to verify that seizures are predictable, one would have to demonstrate substantial evidence that the brain's normal activity differs from the brain's preseizure epileptic activity. If they differ, the next question is "Are different brain's states classifiable?" Thus far, visual inspection of multiple time series of EEG signals in their unprocessed form is still the predominant way of discriminating and classifying EEG patterns in the medical community. This task requires highly trained medical professionals to continuously eyeball the EEG data. None of the current epilepsy studies in the literature is undertaken to develop classification techniques that can be used to systematically classify/differentiate normal and abnormal EEGs. In order to answer these crucial questions, we are motivated to develop such classification technique.

D. Paper Organization

The organization of the succeeding sections of this paper is given as follows. In Section II, the quantification of the brain dynamics and statistical similarity measures of time series for KNN classification are described. The EEG data description and experimental design are presented in Section III. The classification results and the performance characteristics of the proposed technique are addressed in Section IV. The concluding remarks are given in Section V.

II. Method

The classification technique that is developed in this study is comprised of three key features. The first feature is quantitative measures of the brain dynamics. We employ the estimation of the short-term maximum Lyapunov exponent, which has been previously shown to be capable of contemplating dynamical mechanisms of the brain network from EEG signals [20]. The study of the brain dynamics is motivated by the proof concept of chaos theory, i.e., understanding brain dynamics is capable of providing insights about the different states of brain activities that are reflected from pathological dynamical interactions of the brain network [8], [30]. Based on the quantification of the brain dynamics, the subsequent second feature is the statistical similarity measures for similarity/dissimilarity of classifiable features of different brain's physiological states. We herein employ three time series similarity measures to the analysis of brain dynamics in EEG time series. Those measures include *Euclidean, T-Statistical*, and *Dynamic Time Warping (DTW)* distances. The third feature is the classification algorithm, which is based on the KNN rule. Integration of the second and third features results into a novel time series classification technique. We will use this integrated technique to classify normal and preseizure EEGs.

A. Quantification of Brain Dynamics

Quantification of the brain dynamics from EEGs in this study is suitable to the investigation of a nonstationary system such as the brain because it is capable of automatically identifying and appropriately weighing existing transients in the data. This technique is motivated by mathematical models from chaos theory that are used to characterize multidimensional complex systems and reduce the dimensionality of EEGs [31]–[35]. To quantify the brain dynamics, we divide EEG signals into sequential 10.24-s epochs (nonoverlapping windows) to properly account for possible nonstationarities in the epileptic EEG. For each epoch of each channel of the EEG signals, we estimate the measure of chaos, which is known as *short-term maximum Lyapunov exponent*, to quantify the chaoticity of the attractor. A chaotic-system-like human brain is a system in which orbits that originate from similar initial conditions or nearby points in the phase space diverge exponentially in expansion process. The rate of divergence is an important aspect of the dynamical system and is reflected in the value of the Lyapunov exponents. In other words, the Lyapunov exponents measure the average uncertainty along the local eigenvectors of an attractor in the phase space. Next, we will give a short overview of mathematical models that are used in the estimation of the short-term maximum Lyapunov exponent from the EEG signals.

1) EEG Time Series Embedding: In the study of brain dynamics, the initial step in analyzing the dynamical properties of EEG signals is to embed it in a higher dimensional space of dimension p, which enables us to capture the behavior in time of the p variables that are primarily responsible for the dynamics of the EEG. We can now construct p-dimensional vectors $X(t)$, whose components consist of the values of the recorded EEG signal $x(t)$ at p points in time that is separated by a time delay. Construction of the embedding phase space from a data segment $x(t)$ of duration T is made with the method of delays. The vectors X_i in the phase space are constructed as

$$X_i = (x(t_i), x(t_i + \tau), \ldots, x\left(t_i + (p-1)^*\tau\right)) \qquad (1)$$

where τ is the selected time lag between the components of each vector in the phase space, p is the selected dimension of the embedding phase space, and $t_i \in [1, T - (p-1)\tau]$. The vectors X_i in the phase space are illustrated in Fig. 3.

2) Estimation of Short-Term Maximum Lyapunov Exponent $STL_{\max}$: The method for estimation of $STL_{\max}$ for nonstationary data (e.g., EEG time series) is previously explained in [8] and [36]. In this section, we will only give a short description and basic notation of our mathematical models that

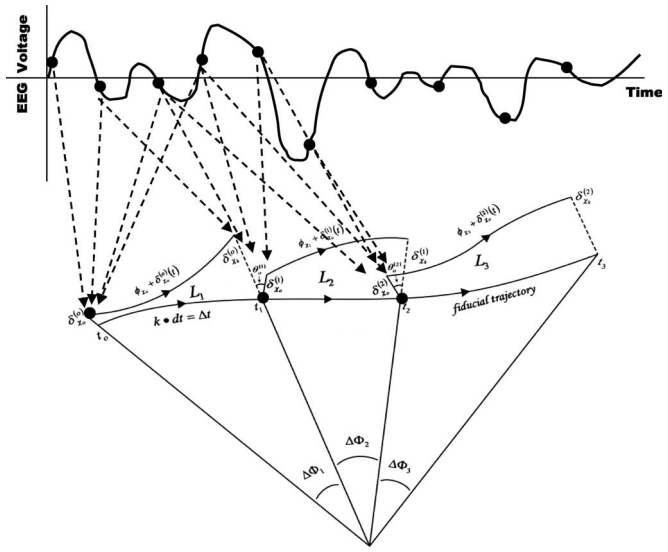

Fig. 3. Diagram illustrating an EEG epoch that is embedded in phase space for the quantification of brain dynamics, assuming that $p = 4$. The fiducial trajectory, i.e., the first three local Lyapunov exponents (L1, L2, L3), is shown.

are used to estimate $STL_{\max}$. First, let us define the notation that is given here.

- N_a is the number of local $STL_{\max}$'s that will be estimated within a duration T data segment. Therefore, if D_t is the sampling period of the time domain data, $T = (N-1)D_t = N_a\Delta t + (p-1)\tau$.
- $X(t_i)$ is the point of the fiducial trajectory $\phi_t(X(t_0))$ with $t = t_i$, $X(t_0) = (x(t_0), \ldots, x(t_0 + (p-1)^*\tau))$, and $X(t_j)$ is a properly chosen vector that is adjacent to $X(t_i)$ in the phase space.
- $\delta X_{i,j}(0) = X(t_i) - X(t_j)$ is the displacement vector at t_i, i.e., a perturbation of the fiducial orbit at t_i, and $\delta X_{i,j}(\Delta t) = X(t_i + \Delta t) - X(t_j + \Delta t)$ is the evolution of this perturbation after time Δt.
- $t_i = t_0 + (i-1)^*\Delta t$ and $t_j = t_0 + (j-1)^*\Delta t$, where $i \in [1, N_a]$ and $j \in [1, N]$ with $j \neq i$.
- Δt is the evolution time for $\delta X_{i,j}$, i.e., the time one allows $\delta X_{i,j}$ to evolve in the phase space. If evolution time Δt is given in seconds, then L is given in bits per second.
- t_0 is the initial time point of the fiducial trajectory and coincides with the time point of the first data in the data segment of analysis. In the estimation of $STL_{\max}$, for a complete scan of the attractor, t_0 should move within $[0, \Delta t]$.
- $STL_{\max}$ is defined as the average of local Lyapunov exponents in the state space and can be calculated by the following equation:

$$STL_{\max} = \frac{1}{N_a \Delta t} \sum_{i=1}^{N_a} \log_2 \frac{|\delta X_{i,j}(\Delta t)|}{|\delta X_{i,j}(0)|}. \qquad (2)$$

B. Similarity Measures

In this study, the choice of similarity measures is a very important step in achieving accurate classification results. To our knowledge, we are the first to develop a framework incorporating statistical time series similarity measures for differentiating EEG signals from normal and preseizure states. We will also apply and compare the performance of those similarity measures, including *Euclidean, T-Statistical (index),* and *DTW* distances. Each of these measures can give us different aspects and insights about the temporal and spatial characteristics of EEG signals. In this section, we will represent the time series of measures of chaos in the notation of $STL_{\max}$ as time series X to simplify the mathematical models.

1) DTW Distance: Given two time series (or vector sequences) X and Y of equal length $|X| = |Y| = n$, pattern similarity is determined by aligning time series X with time series Y with the distortion of alignment $D_{\text{align}}(X, Y)$. DTW is used to compute the best possible alignment warp between two time series by selecting the one with the minimum distortion. In other words, the DTW distance is a distance measure (or similarity measure) between two time series by computing the best possible alignment or the minimum mapping (aligning) distance between two time series. In this study, all our EEG data samples are equal in length; however, the DTW can be extended to the case where the lengths of the two time series are not equal. DTW has been widely used in many contexts including data mining [37], [38], gesture recognition [39], robotics [40], speech processing [41]–[43], and medicine [44].

The problem of calculating the DTW distance can be solved by a dynamic programming approach. The basic concept can be described as follows. First, construct an alignment of every time series vector (data point) in X to match with every corresponding vector in Y. The $n \times n$ alignment matrix will then be constructed by using the Euclidean distance as the local distance between two vectors, i.e., $d(x_i, y_j) = (x_i - y_j)^2$, where the (i, j)th element of the matrix is distance $d(x_i, y_j)$ between the ith point of time series X and the jth point of time series Y. Subsequently, we construct a warp path, i.e., $W = w_1, \ldots, w_K$, starting at the beginning of each time series $w_1 = (1, 1)$ and finishing at the end of both time series $w_K = (n, n)$. Note that K is the length of the warp path and $\max(|X|, |Y|) \leq K < |X| + |Y|$. The kth element of the warp path represents the matching point of two time series, i.e., $w_k = (i, j)$, where (i, j) corresponds to index i from time series X and index j from time series Y (shown in Fig. 4). The warp path can actually be calculated in reverse order, starting at the end of both time series. There is also another constraint on the warp path on the monotonically increasing indices of a warp path. Specifically, if $w_k = (i, j)$ and $w_{k+1} = (i', j')$, then a warp path must satisfy $i \leq i' \leq i + 1$ and $j \leq j' \leq j + 1$.

Note that there can be an exponential number of warping paths that satisfy the preceding conditions. However, the optimal warp path is the one with a minimum warping (distortion) cost defined by

$$D_{\text{align}}(X, Y) = \min \frac{1}{K} \sum_{k=1}^{K} d(w_{ki}, w_{kj}).$$

In a dynamic programming approach, the warp path must either be incremented by one unit (adjacent) or stay at the same i-axis or j-axis. Therefore, we only need to evaluate the

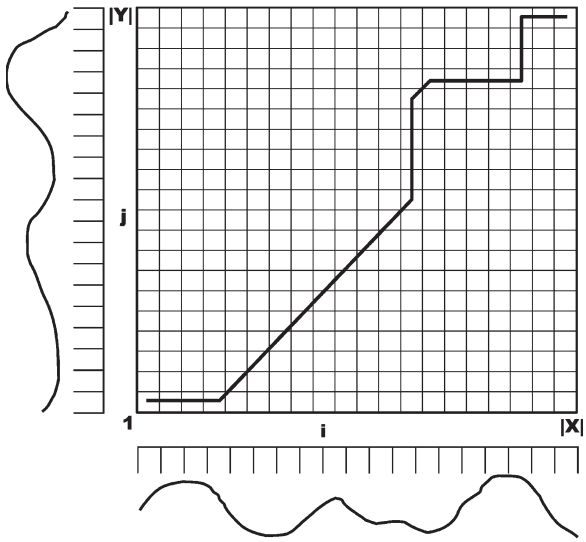

Fig. 4. Warping matrix with the minimum-distance warp path of two time series X and Y.

recurrence of the cumulative distance that is found in the adjacent elements, i.e.,

$$D(i,j) = d(x_i, y_j) + \min \begin{cases} D(i, j-1) \\ D(i-1, j) \\ D(i-1, j-1). \end{cases}$$

This results into the optimality condition for the warping path: $\mathrm{DTW}^*_{xy}(w_k) = d(x_{i(k)}, y_{j(k)}, w_k) r(w_k) + \mathrm{DTW}^*_{xy}(w_{k-1})$.

2) T-Statistical (Index) Distance: The t-statistics is a statistical analysis that is used to examine if two time series are statistically different from each other. In particular, it assesses whether the means of two time series are statistically different from each other. In other words, we commonly use the t-test to determine if two time series differ from each other in a significant way under the assumptions that the paired differences are independent and identically normally distributed. Note that the t-test deals with the problems that are associated with inference based on small samples (epochs of time series), which implies that the calculated mean and standard deviation may deviate from the "real" mean and standard deviation. The t-index is a similarity degree of the t-statistics from the paired t-test for comparisons of means of paired-dependent observations. In this study, we use the t-index as a measure of statistical distance between two time series. Specifically, it is used to estimate the difference of the EEG signals from different brain states. The t-index can be seen as a ratio of the difference between the two means or averages and a measure of the variability or dispersion of the scores. The property of t-index is essentially another example of the signal-to-noise metaphor in research: The difference between the means is the signal, and a measure of variability is essentially noise that may make it harder to see the group difference. The t-index T_{xy} between the time series $X = x_1, \ldots, x_n$ and $Y = y_1, \ldots, y_n$ is then defined as

$$T_{xy} = \frac{\sum_{i=1}^{n} |x_i - y_i|}{\sqrt{n}\sigma_{|x-y|}} \tag{3}$$

where $\sigma_{|x-y|}$ is the sample standard deviation of the absolute difference between time series X and Y estimated over a window with length n. Note that the t-index follows a t-distribution with $n-1$ degrees of freedom.

3) Euclidean Distance: The Euclidean distance is the most commonly used similarity measure. It is easy to understand and bear certain success in many classification problems. It measures the degree of similarity in terms of intensity of the data. In short, the Euclidean distance tells us an average of the difference in intensity of two time series. The Euclidean distance ED between time series X and Y is defined as $ED_{xy} = (\sum_{i=1}^{n}(x_i - y_i)^2)/n$.

C. KNN Classification

KNN is a very intuitive method in which the classifier labels samples (time series) based on their similarity between samples in the training data. In other words, the classifier makes a decision by comparing a new unlabeled sample with the baseline data. This technique represents a bridge between the parametric techniques that require *a priori* knowledge of the distributions underlying the data and nonparametric techniques, which presuppose the functional form of the discriminant surfaces separating the different pattern classes. Applications of KNN have been successful in many areas including handwritten digit recognition [45], gene expression classification [46], and text mining [47]. In general, for a given unlabeled time series X, the KNN rule finds the k "closest" (neighborhood) labeled time series in the training data set and assigns X to the class that appears most frequently in the neighborhood of k time series. Besides the training data, the KNN rule only requires three input parameters that are used for classifying a new unlabeled time series, i.e., the size of the neighborhood k and a similarity function that is used as a measure of "closeness." There are two common rules for classifying the new unlabeled data: 1) majority voting and 2) similarity degree summing. In majority voting, a class (category) gets one vote for each instance of that class in the set of k neighborhood samples. Then, the new data sample is classified to the class with the highest amount of votes. In similarity score summing, each class gets a score that is equal to the summation of the similarity degrees of the instances of that class in the set of k neighborhood samples. Then, the new data sample is classified to the class with the highest similarity degree sum. Note that majority voting is more commonly used because it is less sensitive to outliers. Thus, we shall use majority voting in this study. This is illustrated in Fig. 5, which shows data records, each with two attributes (2-D plot) that are representations of three classes of data (red, blue, and green). In this case, $k = 5$. Thus, the unknown-labeled test point would be labeled by the category of the red points. However, note that we will "train" the classifier to find the best value of k to maximize the classification performance.

III. Materials and Experimental Design

The underlying hypothesis is that the measure of chaos from EEG time series can be used as features to discriminate different stages of the brain dynamics, and these features from the

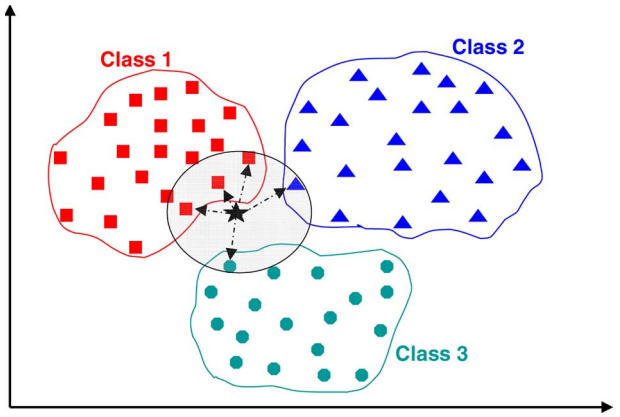

Fig. 5. KNN query starts at the test point and grows a spherical region until it encloses k training samples, and it labels the test point by a majority voting of these samples. In this case where $k = 5$, the test point would be labeled by the category of the red points.

TABLE I
EEG DATASET CHARACTERISTICS

Patient ID	Seizure types	Duration of EEG (days)	♯ of seizures
1	CP, SC	3.55	7
2	CP, GTC, SC	10.93	7
3	CP	8.85	22
4	CP, SC	5.93	19
5	CP, SC	13.13	17
6	CP, SC	11.95	17
7	CP, SC	3.11	9
8	CP, SC	6.09	23
9	CP, SC	11.53	20
10	CP	9.65	12
Total		84.71 days	153

CP-Complex Partial; SC-Subclinical; GTC-Generalized Tonic/Clonic

same state should be more similar than the one from different states. In other words, the characteristics of the brain dynamics during the normal state should be more similar to each other than those during the preseizure state and vice versa. This study is undertaken to demonstrate that the classification technique that was described in the previous section can differentiate the EEG signals from the normal and preseizure states. The classifiability of different brain states will be demonstrated through the quantitative analysis of the brain dynamics $STL_{\max}$ by using the *time series KNN classification* that is proposed here.

A. EEG Data Acquisition

The data sets consisted of continuous long-term (3–13 days) multichannel intracranial EEG recordings from bilaterally surgically implanted macroelectrodes in the hippocampus, temporal, and frontal lobe cortexes of ten epileptic patients with medically intractable temporal lobe epilepsy (outlined in Table I). The recordings were obtained as part of a presurgical clinical evaluation, using a Nicolet BMSI 4000 recording system with amplifiers of an input range of 0.6 mV, sampling rate of 200 Hz, and filters with a frequency range of 0.5–70 Hz. Each recording included a total of 26–32 intracranial electrodes (eight subdural and six hippocampal depth electrodes for each cerebral hemisphere, and a strip of four additional electrodes if deemed necessary by the neurologist). The recorded EEG signals were digitized and stored on magnetic media for subsequent offline analysis. These EEG recordings were viewed by two independent electroencephalographers to determine the number and type of recorded seizures, seizure onset and end times, and seizure onset zones.

B. Data Selection and Sampling Procedure

In this study, the classification will be performed separately for each subject. For consistency, we analyze and investigate EEG time series only from 26 standard electrodes from every patient. First, we randomly sample two groups (normal and preseizure) of 5-min EEG epochs from the continuous recordings in each patient. Per seizure, five EEG epochs from each of the normal and preseizure states are randomly and uniformly sampled. Normal EEG samples are selected from EEG recordings that are more than 8 h apart from a seizure. Preseizure EEG epochs are selected from EEG recordings during the 30-min interval before. For example, patient 1 had seven seizures; therefore, 70 EEG epochs (35 normal and 35 preseizure) will be sampled from patient 1's data set.

C. Classification Procedure

After two groups of EEG epochs are collected, we first calculate the measures of chaos ($STL_{\max}$ profiles) from EEG signals using the method that was described in the previous section. Each measure was calculated continuously for each nonoverlapping 10.24-s segment of EEG data. After we calculate an $STL_{\max}$ profile for each EEG channel, we then perform the KNN classification by categorizing an unknown sample using the KNN rules based on the average statistical distance over 26 electrodes. There are a few parameters in the KNN classification that need to be trained. This can be done by dividing the EEG data into a training set (baseline) and a test set. There are many alternatives on how to divide the data into the training and test sets. We describe the technique that is used to perform this operation in the next section. For each training set, we use different k values to train for the optimal k value that gives the "best" result in classification. Then, we use EEG epochs from the test set to evaluate the performance of this k value. In this study, we will vary the value of k from 3 to 13 (only odd numbers). The optimal k value will be determined in the training phase (discussed in next section). Note that we apply the KNN using all three time series similarity measures that are discussed in this paper. Note that, although the Euclidean distance is commonly used in a standard KNN algorithm, the KNN is very sensitive to the choice of similarity measure that is used, and the EEG characteristics are much more complicated than the Euclidean distance can capture. We anticipate that more sophisticated similarity measures such as t-index and DTW will outperform the standard Euclidean distance.

D. Training and Testing: Cross Validation

There are many choices of how to divide the data into training and test sets. In order to reduce the bias of training and test data, we propose to employ an n-fold cross-validation technique

		Predict: Abnormal	Predict: Normal
Actual	Abnormal	True Positive	False Negative
	Normal	False Positive	True Negative

Fig. 6. Evaluation concept of classification results. Note that we define "preseizure" as abnormal.

to train for the best value of "k" in this study. The cross-validation technique is extensively used as a method to estimate the generalization error based on "resampling." In other words, these techniques will be implemented during the training phases to estimate how well the classifiers learning from the training EEG data is going to perform on future as-yet-unseen data in the testing phase. Generally, with n-fold cross validation, EEGs data will be divided into n subsets of (approximately) equal size. The proposed KNN will be trained and tested n times, in which one of the subsets from training is left out each time, and tested on the omitted subset [22], [48].

As mentioned in [49], the result from one n-fold cross validation may not be reliable. In order to have low mean square error and bias, we perform ten repetitions of cross validations. After ten cross-validation replications for each value of "n" (folds), we compare the average of the classification performance on different values of "k" (nearest neighbors). In this study, we select the values of n to be 3 and 5 as the EEG sample sizes in some patients are very small. It is important to note that the classification techniques will be trained and tested individually for each subject.

E. Performance Evaluation of Classification Schemes

In general, to evaluate the classifier, we categorize the classification into two classes: 1) positive (preseizure) and 2) negative (normal). Then, we consider four subsets of classification results: 1) true positives (TPs), which denote correct classifications of positive cases; 2) true negatives (TNs), which denote correct classifications of negative cases; 3) false positives (FPs), which denote incorrect classifications of negative cases into class positive; and 4) false negatives (FNs), which denote incorrect classifications of positive cases into class negative. To better explain the concept of the evaluation of classifiers, let us consider in the case of the detection of preseizure EEG data (see Fig. 6).

Sensitivity and specificity are widely used in the medical domain as classification performance measures. Sensitivity measures the fraction of positive cases that are classified as positive. Specificity measures the fraction of negative cases that are classified as negative. In fact, the sensitivity can be considered as a probability of accurately classifying EEG samples in the preseizure case. The specificity can be considered as a probability of accurately classifying EEG samples in the normal case. The sensitivity and specificity are defined as follows:

$$\text{sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}.$$

F. Selection of the Best Classification Schemes

To quantitatively select the best classification scheme as well as the best parameter setting (i.e., the best k value and similarity measure), we will implement a statistical method that is called receiver operating characteristics (ROCs) and derived from the detection theory to identify the optimal parameter setting. The ROC analysis is used to indicate an appropriate tradeoff that one can achieve between the detection rate (sensitivity, which is plotted on the Y-axis) that is desired to be maximized and the false alarm rate ($1 -$ specificity, which is plotted on the X-axis) that is desirable to be minimized. Basically, the sensitivity and the specificity of each classification scheme for each parameter setting are calculated. By definition of ROC analysis, the best scheme is selected such that it is closest to the ideal classifier (best performance). For each individual patient, the best classification scheme can be identified by selecting the scheme whose performance is closest to the ideal point (sensitivity $= 1$ and $1 -$ specificity $= 0$), i.e., the scheme that is closest to the top left-hand corner on the ROC plot will be selected. The performance evaluation of the classification schemes and the results for each of classification schemes will be discussed in more detail in the next sections.

IV. Result

Among all ten patients, we sample and preprocess the total of 1530 EEG epochs from the normal and preseizure states, as described in the previous section. Each epoch contains 780 data points (26 electrodes, 5-min epoch $=$ 30 points). The next step is to train our classifier for the best value of k-nearest neighbors. In this experiment, we try different k values ranging from 3 to 13 (note that k should be an odd number). As mentioned earlier, threefold and fivefold cross validations are employed in this study. We will test the KNN algorithm using three time series similarity measures (i.e., Euclidean, t-index, and DTW). The sensitivity and specificity of each parameter setting (value of k and choice of similarity measure) are calculated. It is worth noting that we are studying two-class classification in this study. Thus, the specificity of classifying preseizure EEGs is equivalent to the sensitivity of classifying normal EEGs. In this section, we will use these terms interchangeably. Finally, based on the classification performances from all combinations of parameter settings, the best parameter setting for each patient is selected from the ROC analysis.

A. ROC Analysis

In order to achieve the most appropriate tradeoff to maximize the detection rate (sensitivity, which is plotted on the Y-axis) and minimize the false alarm rate ($1 -$ specificity, which is plotted on the X-axis), the best classification scheme for each patient is selected such that it is closest to the ideal classifier (best possible performance) at the top left-hand corner of the ROC plot. Per individual patient, we apply ROC analysis to indicate the best parameter setting in the KNN classification. Fig. 7 illustrates an ROC plot of the KNN classification using DTW applied to patient 10 using fivefold cross validation with

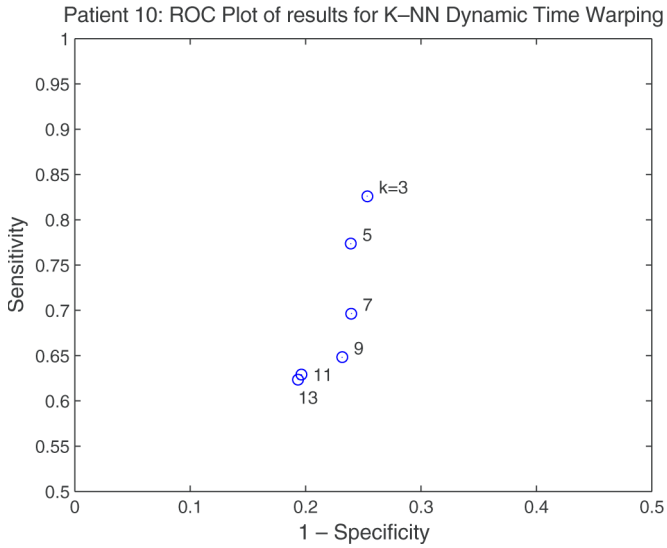

Fig. 7. ROC plot for DTW in patient 10 using fivefold cross validation to train and test the algorithm.

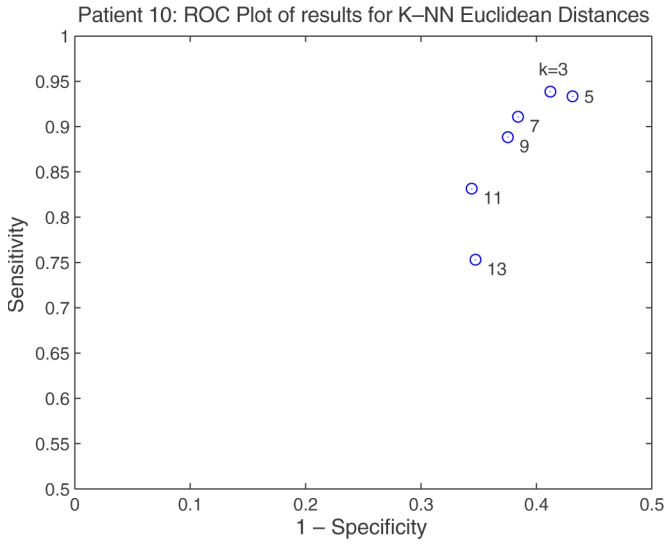

Fig. 8. ROC plot for EDs in patient 10 using fivefold cross validation to train and test the algorithm.

ten replications to train and test the KNN algorithm. Note that each point in the plot represents a value of k-nearest neighbors (ranging from 3 to 13). In Fig. 7, we observe that the KNN classification with DTW achieves its best performance when $k = 3$, which yields a sensitivity of about 84% and a specificity of about 75%. Figs. 8 and 9 illustrate the ROC plots of the KNN classification using *Euclidean* and *T-Statistical* distances, respectively, in patient 10 using fivefold cross validation with ten replications. In Fig. 8, we observe that the KNN classification with Euclidean distance achieves its best performance when $k = 11$, which yields a sensitivity of about 84% and a specificity of about 65%. In Fig. 9, we observe that the KNN classification with T-Statistical distance achieves its best performance when $k = 3$, which yields a sensitivity of about 89% and a specificity of about 91%. From these ROC plots, in the case of patient 10, it is clear that the KNN classification

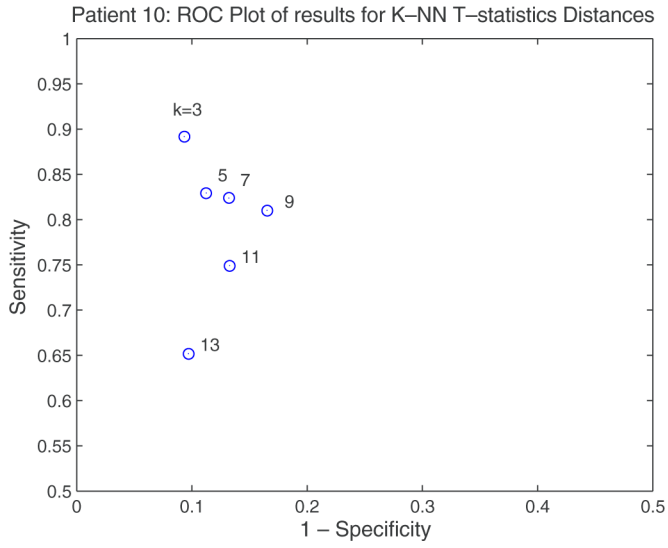

Fig. 9. ROC plot for t-statistics in patient 10 using fivefold cross validation to train and test the algorithm.

using the T-Statistical distance outperforms all other time series similarity measures.

B. *Classification Results*

We apply the ROC analysis, as described in the previous section, to identify the best classification scheme (value of k and choice of similarity measure) per individual patient. The performance characteristics of the best scheme for each individual patient in threefold and fivefold cross validations are listed in Table II. From the table, we can see the consistent result of the best performance from the T-Statistical distance. The results follow the intuitive sense in the classification theory that the performance on fivefold cross validation should be better the one on threefold cross validation, with the reason being that the training sample size is larger. It is clear that the KNN classification is sensitive to the size of training samples, and we expect the performance to be better with a higher value of n-fold. With fivefold cross validation, the KNN classification achieved about 81% sensitivity and 73% specificity, on average, over ten patients. With threefold cross validation, the performance is slightly worse as the KNN classification achieved about 79% sensitivity and 71% specificity, on average, over 10 patients. As observed in Table II, the optimal value of k-nearest neighbors for the T-Statistical distance ranges from 3 to 13. While $k = 3$ tends to be the best global setting, this implies that the T-Statistical distance is capable of capturing specific patterns that are manifested in EEG time series at some specific location at a specific time. This will lead to a computational complexity advantage. As the value of k is lower, the algorithm requires less computational time to search for the nearest neighbors. Thus, the algorithm will run much faster.

Figs. 10 and 11 illustrate the overall classification results that are tested in ten patients with optimal scheme (k value and statistical distance) per individual patient to train and test the algorithm using fivefold cross validation and threefold cross validation, respectively. In Fig. 10, employing the fivefold cross-validation technique, with the optimal k value using the

TABLE II
PERFORMANCE CHARACTERISTICS OF THE BEST CLASSIFICATION SCHEME FOR EACH INDIVIDUAL PATIENT USING THREEFOLD AND FIVEFOLD CROSS VALIDATION ON ONLY $STL_{\max}$ VALUES

Patient	5-fold Cross Validation			3-fold Cross Validation		
	Sensitivity	Specificity	Best Setting	Sensitivity	Specificity	Best Setting
1	92.30%	56.70%	k= 3, T-index	85.71%	49.52%	k= 3, T-index
2	87.90%	69.50%	k= 3, T-index	87.62%	68.57%	k= 3, T-index
3	76.77%	84.96%	k= 3, T-index	77.43%	82.73%	k= 5, T-index
4	86.48%	57.00%	k=11, T-index	82.63%	59.30%	k=13, T-index
5	72.96%	80.44%	k= 3, T-index	71.77%	78.43%	k= 3, T-index
6	82.13%	67.62%	k= 3, T-index	73.73%	65.30%	k= 3, T-index
7	82.47%	87.13%	k= 3, T-index	82.96%	85.19%	k= 3, T-index
8	84.88%	74.46%	k= 3, T-index	83.48%	73.19%	k= 3, T-index
9	72.17%	60.50%	k= 3, T-index	69.50%	56.17%	k= 3, T-index
10	89.18%	90.64%	k= 3, T-index	84.17%	89.72%	k= 3, T-index
Average	81.29%	72.86%		78.61%	71.02%	

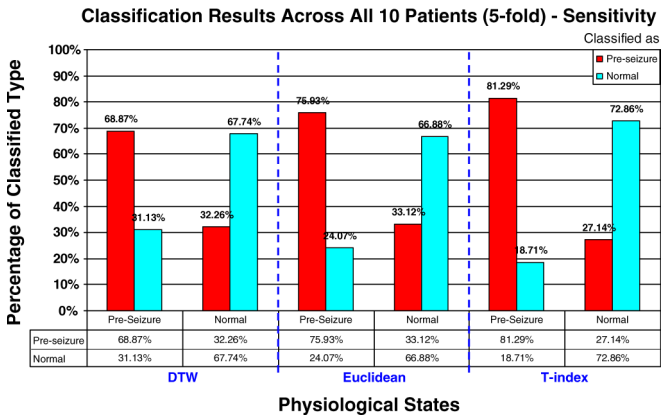

Fig. 10. Overall classification results among all similarity measures using fivefold cross validation to train and test the algorithm.

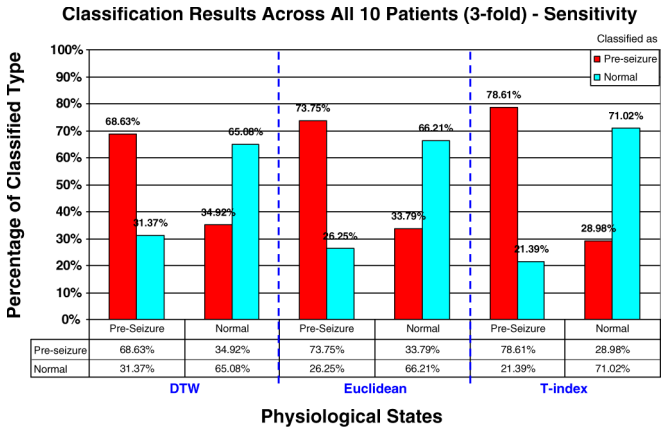

Fig. 11. Overall classification results among all similarity measures using threefold cross validation to train and test the algorithm.

T-Statistical distance, the KNN classification can achieve a sensitivity and specificity of about 81% and 73%, respectively. With the optimal k value using the DTW distance, the KNN classification can achieve a sensitivity and specificity of about 68% and 66%, respectively. With the optimal k value using the Euclidean distance, the KNN classification can achieve a sensitivity and specificity of about 75% and 67%, respectively. In Fig. 11, employing the threefold cross-validation technique, with the optimal k value using the T-Statistical distance, the KNN classification can achieve a sensitivity and specificity of about 79% and 71%, respectively. With the optimal k value using the DTW distance, the KNN classification can achieve a sensitivity and specificity of about 68% and 65%, respectively. With the optimal k value using the Euclidean distance, the KNN classification can achieve a sensitivity and specificity of about 73% and 66%, respectively.

It can also be observed that the KNN classification achieved lower specificity than all other cases. One speculation is that there are more variabilities in the EEG patterns during the normal period than during the preseizure period. In other words, because the preseizure EEG patterns are more similar as they evolve in the same state transition (seizure development), the KNN classification is able to better capture similarity in preseizure patterns.

Note that the results in this study are consistent with the results from our previous study in classifying EEGs [22]. Surprisingly, the use of the DTW distance did not improve the performance of the KNN classification to achieve better classification results. Nevertheless, with a choice of any similarity measures in general, the KNN algorithm shows a potential to be a good time series classification as it preserves both the spatial and temporal properties of the data. These results confirm our hypothesis that the brain's states are classifiable based on quantitative analyses of EEG. The framework of the classifiers that are proposed in this study can be extended to the development of an abnormal brain activity classifier, which will be described in the next section.

V. CONCLUSION AND DISCUSSION

This study addresses the open question of the classifiability of the brain's preseizure and normal states. The experimental results on the performance of the *time series* KNN are very encouraging. These results represent a proof concept of the successful brain dynamics quantification through EEGs. This concept can provide insights and characterize different states of brain activities that are reflected from pathological dynamical interactions of the brain network. In addition, these results also confirm our hypothesis that it is possible to differentiate and classify the brain's preseizure and normal activities based on optimization, data mining, and dynamical system approaches in multichannel intracranial EEG recordings. The reliable classification is conceivable because, for the vast majority of seizures, the spatiotemporal dynamical features of the preseizure state

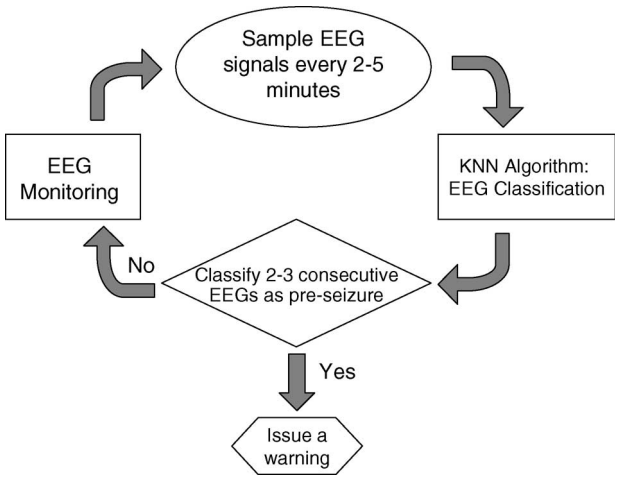

Fig. 12. Flowchart of an automated system, which can be used as a base module of ABAC or ASPS.

sufficiently differ from those of the normal state. It can be easily observed that the brain dynamics from different states of the same patient are qualitatively and quantitatively differentiable. In addition, we notice that the performance of the KNN algorithm is irrelevant to the characteristic of the patients. Surprisingly, a choice of smaller k value for all time series similarity measures tends to be favorable. In addition, the proposed technique is very fast and scalable, and it does not require any assumption of distribution of the EEG data. All of the programming was done in Matlab environment on a desktop computer with Pentium Xeon 2.8 GHz with 2 GB of random access memory. The running time for the statistical cross-validation technique is less than 5 min on average. The bottleneck of the algorithm is twofold. One is the calculation of the measure of chaos, and another is the calculation of the DTW and T-Statistical distances. Calculating the DTW distance, we are required to solve multiple dynamic programs in addition to the calculation of the Euclidean distance. Calculating the T-Statistical distance, we are required to estimate the standard deviation of the difference of two time series. Both of these steps can be improved when we implement an online real-time analysis in our future study.

The results of this study provide a framework of methodology that can be used to elucidate the dynamical mechanisms of the brain mechanisms underlying the epileptogenic process. In addition, the results can be extended to the development of an online brain monitoring system. For instance, the classification algorithm can be embedded into the EEG monitoring unit. The unit can collect normal and preseizure EEG samples as a baseline database and subsequently perform the EEG classification by periodically sampling the ongoing EEG signals. If the EEG samples are classified as preseizure, one can trigger the warning of an impending seizure. Success of this study is a necessary first step in the development of an abnormal brain activity classifier (ABAC) and an automated seizure prediction system (ASPS), which are computer-based systems that are used to predict seizures or classify abnormal EEGs (see, e.g., Fig. 12). In order to make another step closer to the ABAC, in our future study, we would like to extend the proposed classification technique to scalp EEG data as it is more practical and feasible to obtain scalp EEG data. In this study, the intracranial EEG data that are studied here have good characteristics for classification because they are clean of artifacts that are caused by the EEG machines as well as the patients' movements (e.g., coughing and swallowing). As opposed to intracranial EEGs, the most challenging issue is that scalp EEGs are much more sensitive and susceptible to artifacts. These will make it very hard for our algorithm to classify the real physiological states of the brain. An analysis of our algorithm on scalp EEGs will be studied in the future. Such classification algorithms for scalp EEGs can also be used to develop a quick screening procedure to classify normal and epilepsy patients. In addition, the feature selection study will be possible in the future. This study will help us to select electrodes that show prominent changes and give us the best classification results, which might lead us to the solution to the epileptogenic zone localization problem.

Another prospective of this study is to understand exactly how biology makes certain people vulnerable to seizures. The reigning theory holds that abnormal activities/scenarios affect the risk of seizures being developed. Success of this project will help neurologists to find better faster ways of gauging the risk of seizures being developed in epileptic patients. We believe that the lessons that were learned in controlling abnormal heart rhythms could be applied to controlling brain arrhythmias. Traditionally, cardiac arrhythmias were treated with medicines. Today, many are treated with automatic defibrillators, which are devices that are surgically placed in patients' chests. The devices monitor heart rhythms, and when the rhythms become abnormal, they will give a very small jolt of electricity to put the heart back in its normal rhythm. In the same view, epilepsy may respond to a similar kind of closed-loop feedback warning system. The device would predict the onset of a seizure and prevent the seizure before it occurs. These results are a very important proof of concept investigation of the development of implantable seizure feedback control devices to directly regulate therapeutic pharmacological or physiological intervention to prevent impending seizures or other brain disorders. For example, such an intervention might be achieved by electrical or magnetic stimulation (e.g., vagal nerve stimulation) or by a timely release of an anticonvulsant drug. In the future, a new outlook for epilepsy treatment would be to trigger a pulse intervention long (ranging from minutes to hours) before the seizure actually begins. It is possible that an electrical stimulus that is delivered during this preseizure period could abort an impending seizure before it actually occurs. A warning of several minutes in duration could even trigger the release of an anticonvulsant drug in time to abort an impending seizure. The outcome of this study will form a bridge between seizure prediction research and the implementation of biofeedback seizure control devices, which will be a revolutionary approach for handling epileptic seizures (i.e., **the brain pacemaker**).

ACKNOWLEDGMENT

The authors would like to thank Prof. E. Keogh, Prof. P. M. Pardalos, Prof. L. D. Iasemidis, and Prof. J. C. Sackellares for their fruitful comments and discussion.

REFERENCES

[1] O. Cockerell, I. Eckle, D. Goodridge, J. Sander, and S. Shorvon, "Epilepsy in a population of 6000 re-examined: Secular trends in first attendance rates, prevalence, and prognosis," *J. Neurol., Neurosurg., Psychiatr.*, vol. 58, no. 5, pp. 570–576, 1996.

[2] W. H. O. (WHO), *Epilepsy: Historical Overview*, 2004. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs168/en/

[3] C. Begley, M. Famulari, J. Annegers, D. Lairson, T. Reynolds, S. Coan, S. Dubinsky, M. Newmark, C. Leibson, E. So, and W. Rocca, "The cost of epilepsy in the United States: An estimate from population-based clinical and survey data," *Epilepsia*, vol. 41, no. 3, pp. 342–351, 2000.

[4] C. Begley, J. Annegers, D. Lairson, T. Reynolds, and W. Hauser, "Cost of epilepsy in the United States: A model based on incidence and prognosis," *Epilepsia*, vol. 35, no. 6, pp. 1230–1243, 1994.

[5] L. Iasemidis, "Epileptic seizure prediction and control," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 5, pp. 549–558, May 2003.

[6] K. Lehnertz and B. Litt, "The first international collaborative workshop on seizure prediction: Summary and data description," *J. Clin. Neurophysiol.*, vol. 116, no. 3, pp. 493–505, 2005.

[7] B. Litt and J. Echauz, "Prediction of epileptic seizures," *Lancet Neurol.*, vol. 1, no. 1, pp. 22–30, May 2002.

[8] L. Iasemidis, "On the dynamics of the human brain in temporal lobe epilepsy," Ph.D. dissertation, Univ. Michigan, Ann Arbor, MI, 1991.

[9] P. Pardalos, W. Chaovalitwongse, L. Iasemidis, J. Sackellares, D.-S. Shiau, P. Carney, O. Prokopyev, and V. Yatsenko, "Seizure warning algorithm based on spatiotemporal dynamics of intracranial EEG," *Math. Program.*, vol. 101, no. 2, pp. 365–385, 2004.

[10] C. Elger and K. Lehnertz, "Seizure prediction by non-linear time series analysis of brain electrical activity," *Eur. J. Neurosci.*, vol. 10, no. 2, pp. 786–789, Feb. 1998.

[11] K. Lehnertz and C. Elger, "Can epileptic seizures be predicted? Evidence from nonlinear time series analysis of brain electrical activity," *Phys. Rev. Lett.*, vol. 80, no. 22, pp. 5019–5022, Jun. 1998.

[12] F. Mormann, T. Kreuz, C. Rieke, R. Andrzejak, A. Kraskov, P. David, C. Elger, and K. Lehnertz, "On the predictability of epileptic seizures," *J. Clin. Neurophysiol.*, vol. 116, no. 3, pp. 569–587, 2005.

[13] M. L. V. Quyen, J. Martinerie, M. Baulac, and F. Varela, "Anticipating epileptic seizures in real time by non-linear analysis of similarity between EEG recordings," *NeuroReport*, vol. 10, no. 10, pp. 2149–2155, Jul. 1999.

[14] J. Martinerie, C. V. Adam, and M. L. V. Quyen, "Epileptic seizures can be anticipated by non-linear analysis," *Nat. Med.*, vol. 4, no. 10, pp. 1173–1176, Oct. 1998.

[15] M. Harrison, M. Frei, and I. Osorio, "Accumulated energy revisited," *J. Clin. Neurophysiol.*, vol. 116, no. 3, pp. 527–531, 2005.

[16] B. Litt, R. Esteller, J. Echauz, D. Maryann, R. Shor, T. Henry, P. Pennell, C. Epstein, R. Bakay, M. Dichter, and G. Vachtservanos, "Epileptic seizures may begin hours in advance of clinical onset: A report of five patients," *Neuron*, vol. 30, no. 1, pp. 51–64, Apr. 2001.

[17] K. Jerger, S. Weinstein, T. Sauer, and S. Schiff, "Multivariate linear discrimination of seizures," *J. Clin. Neurophysiol.*, vol. 116, no. 3, pp. 545–551, 2005.

[18] W. Chaovalitwongse, L. Iasemidis, P. Pardalos, P. Carney, D.-S. Shiau, and J. Sackellares, "Performance of a seizure warning algorithm based on the dynamics of intracranial EEG," *Epilepsy Res.*, vol. 64, no. 3, pp. 93–133, May 2005.

[19] W. Chaovalitwongse, P. Pardalos, L. Iasemidis, J. Sackellares, and D.-S. Shiau, "Optimization of spatio-temporal pattern processing for seizure warning and prediction," U.S. Patent application filed August 2004, Attorney Docket No. 028724-150, 2004.

[20] W. Chaovalitwongse, P. Pardalos, L. Iasemidis, D.-S. Shiau, and J. Sackellares, "Applications of global optimization and dynamical systems to prediction of epileptic seizures," in *Quantitative Neuroscience*, P. Pardalos, J. Sackellares, L. Iasemidis, and P. Carney, Eds. Norwell, MA: Kluwer, 2003, pp. 1–36.

[21] W. Chaovalitwongse, P. Pardalos, L. Iasemidis, D. Shiau, and J. Sackellares, "Dynamical approaches and multi-quadratic integer programming for seizure prediction," *Optim. Methods Softw.*, vol. 20, no. 2/3, pp. 383–394, 2005.

[22] W. Chaovalitwongse, P. Pardalos, and O. Prokoyev, "Electroencephalogram (EEG) time series classification: Applications in epilepsy," *Ann. Oper. Res.*, vol. 148, no. 1, pp. 227–250, Nov. 2006.

[23] R. Esteller, J. Echauz, M. D'Alessandro, G. Worrell, S. Cranstoun, G. Vachtsevanos, and B. Litt, "Continuous energy variation during the seizure cycle: Towards an on-line accumulated energy," *J. Clin. Neurophysiol.*, vol. 116, no. 3, pp. 517–526, 2005.

[24] W. Chaovalitwongse, P. Pardalos, and O. Prokoyev, "Reduction of multi-quadratic 0–1 programming problems to linear mixed 0–1 programming problems," *Oper. Res. Lett.*, vol. 32, no. 6, pp. 517–522, 2004.

[25] W. Chaovalitwongse, "Optimization and dynamical approaches in nonlinear time series analysis with applications in bioengineering," Ph.D. dissertation, Univ. Florida, Gainesville, FL, 2003.

[26] P. Pardalos and J. Principe, *Biocomputing*. Norwell, MA: Kluwer, 2003.

[27] L. Iasemidis, D.-S. Shiau, W. Chaovalitwongse, J. Sackellares, P. Pardalos, P. Carney, J. Principe, A. Prasad, B. Veeramani, and K. Tsakalis, "Adaptive epileptic seizure prediction system," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 5, pp. 616–627, May 2003.

[28] J. Sackellares, L. Iasemidis, D.-S. Shiau, L. Dance, P. Pardalos, and W. Chaovalitwongse, "Optimization of multi-dimensional time series processing for seizure warning and prediction," International Patent Application filed August 2003, Attorney Docket No. 028724-142, 2003.

[29] J. Sackellares, L. Iasemidis, V. Yatsenko, D.-S. Shiau, P. Pardalos, and W. Chaovalitwongse, "Multi-dimensional multi-parameter time series processing for seizure warning and prediction," International Patent Application filed Septempber 2003, Attorney Docket No. 028724-143, 2003.

[30] L. Iasemidis and J. Sackellares, "The evolution with time of the spatial distribution of the largest Lyapunov exponent on the human epileptic cortex," in *Measuring Chaos in the Human Brain*, D. Duke and W. Pritchard, Eds. Singapore: World Scientific, 1991, pp. 49–82.

[31] A. Babloyantz and A. Destexhe, "Low dimensional chaos in an instance of epilepsy," *Proc. Nat. Acad. Sci. USA*, vol. 83, no. 10, pp. 3513–3517, May 1986.

[32] L. Iasemidis, H. Zaveri, J. Sackellares, and W. Williams, "Phase space analysis of EEG in temporal lobe epilepsy," in *Proc. 10th Annu. Int. Conf. IEEE Eng. Med. and Biol. Soc.*, 1988, pp. 1201–1203.

[33] N. Packard, J. Crutchfield, and J. Farmer, "Geometry from time series," *Phys. Rev. Lett.*, vol. 45, no. 9, pp. 712–716, Sep. 1980.

[34] P. Rapp, I. Zimmerman, and A. M. Albano, "Experimental studies of chaotic neural behavior: Cellular activity and electroencephalographic signals," in *Nonlinear Oscillations in Biology and Chemistry*, H. Othmer, Ed. New York: Springer-Verlag, 1986, pp. 175–205.

[35] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence.* Lecture Notes in Mathematics, D. Rand and L. Young, Eds. New York: Springer-Verlag, 1981.

[36] A. Wolf, J. Swift, H. Swinney, and J. Vastano, "Determining Lyapunov exponents from a time series," *Phys. D*, vol. 16, no. 3, pp. 285–317, 1985.

[37] E. Keogh and M. Pazzani, "Scaling up dynamic time warping for datamining applications," in *Proc. 6th ACM SIGKDD Int. Conf. Knowl. Discov. and Data Mining*, 2000, pp. 285–289.

[38] D. Berndt and J. Clifford, "Using dynamic time warping to find patterns in time series," in *Proc. AAAI Workshop KDD*, 1994, pp. 359–370.

[39] D. Gavrila and L. Davis, "Towards 3-D model-based tracking and recognition of human movement: A multi-view approach," in *Proc. Int. Workshop Autom. Face- and Gesture-Recog.*, 1995, pp. 272–277.

[40] M. Schmill, T. Oates, and P. Cohen, "Learned models for continuous planning," in *Proc. 7th Int. Workshop Artif. Intell. and Statist.*, 1999, pp. 278–282.

[41] L. Rabiner and B. Juang, *Fundamentals of Speech Recognition*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[42] H. Shimodaira, K. I. Noma, M. Naka, and S. Sagayama, "Support vector machine with dynamic time-alignment kernel for speech recognition," in *Proc. Eurospeech*, 2001, pp. 1841–1844.

[43] V. Wan and J. Carmichael, "Polynomial dynamic time warping kernel support vector machines for dysarthric speech recognition with sparse training data," in *Proc. Interspeech*, 2005, pp. 3321–3324.

[44] E. Caiani, A. Porta, G. Baselli, M. Turiel, S. Muzzupappa, F. Pieruzzi, C. Crema, A. Malliani, and S. Cerutti, "Warped-average template technique to track on a cycle-by-cycle basis the cardiac filling phases on left ventricular volume," *IEEE Comput. Cardiol.*, vol. 25, no. 98, CH36292, 1998.

[45] Y. Lee, "Handwritten digit recognition using k nearest neighbour, radial basis function, and back-propagation neural networks," *Neural Comput.*, vol. 3, no. 3, pp. 440–449, 1991.

[46] T. Golub, D. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J. P. Mesirov, H. Coller, M. L. Loh, J. Downing, M. Caligiuri, C. Bloomeld, and E. Lander, "Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring," *Science*, vol. 286, no. 5439, pp. 531–537, Oct. 1999.

[47] S. M. Weiss, C. Apte, F. J. Damerau, D. E. Johnson, F. J. Oles, T. Goetz, and T. Hampp, "Maximizing text-mining performance," *IEEE Intell. Syst.*, vol. 14, no. 4, pp. 63–69, Jul./Aug. 1999.

[48] B. Efron, "Estimating the error rate of a prediction rule: Improvement on cross-validation," *J. Amer. Stat. Assoc.*, vol. 78, no. 382, pp. 316–331, Jun. 1983.
[49] A. Molinaro, R. Simon, and R. Pfeiffer, "Prediction error estimation: A comparison of resampling methods," *Bioinformatics*, vol. 21, no. 15, pp. 3301–3307, 2005.

Wanpracha Art Chaovalitwongse (M'05) received the B.S. degree in telecommunication engineering from the King Mongkut Institute of Technology Ladkrabang, Bangkok, Thailand, in 1999 and the M.S. and Ph.D. degrees in industrial and systems engineering from the University of Florida, Gainesville, in 2000 and 2003, respectively.

He is currently an Assistant Professor of industrial and systems engineering in the Department of Industrial and Systems Engineering, Rutgers University, Piscataway, NJ, where he does research on optimization, data mining, and statistical analysis in medical problems, multicast routing, and supply chain management. In 2003, he was with the NIH-funded Brain Dynamics Laboratory, Brain Institute, and the Departments of Neuroscience and Industrial and Systems Engineering, University of Florida. Before joining Rutgers University, he was with the Corporate Strategic Research, ExxonMobil Research & Engineering in 2004, where he managed research on developing efficient mathematical models and novel statistical data analyses for upstream and downstream business operations.

Ya-Ju Fan (S'07) received the B.B.A. degree in production and operations management from Fu Jen Catholic University, Taipei, Taiwan, R.O.C., and the M.S. degree in industrial engineering from the University of Wisconsin–Madison, Madison. She is currently working toward the Ph.D. degree at Rutgers University, Piscataway, NJ.

Working with Prof. W. Art Chaovalitwongse, she is engaged in research on optimization and data mining.

Rajesh C. Sachdeo received the B.S. degree in premedicine from Ewing Christian College, Allahabad, India, in 1969 and the M.D. degree from Christian Medical College, Vellore, India in 1973. He had his residency training at V.A. Hines and Loyola University Medical Center, Maywood, IL, in 1979 and his fellowship training in neurophysiology and epilepsy from Rush-Presbyterian St. Luke's Medical Center, Chicago, IL, in 1980.

He is currently a Clinical Professor in the Department of Pediatrics, Jersey Shore University Medical Center, Neptune, NJ. He was a Clinical Professor of neurology at St. Peter's University Hospital, New Brunswick, NJ, and the Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, Piscataway, NJ. He is also the Director of the New Jersey Comprehensive Epilepsy Center. Before joining St. Peter's, he was an Assistant Professor in medicine at the University of South Florida, Tampa, an Assistant Chief of neurology at V.A. Medical Center, Lyons, NJ, and an Adjunct Faculty of pharmacy and administration at Rutgers University, Piscataway.

Dr. Sachdeo is a Fellow of the American Academy of Neurology, American EEG Society, American Epilepsy Society, and Academy of Medicine; an honorary member of the New Jersey Academy of Developmental Medicine; and a member of the American Association of Electromyography and Electrodiagnosis, American Academy of Clinical Neurophysiology, and Southern Neurological Society. He has received numerous awards.