A Novel Probabilistic Framework to Personalize Online Epileptic Seizure Prediction *

Shouyi Wang
University of Texas at Arlington
500 West First Street
Arlington, Texas 76019
shouyiw@uta.edu

Stephen Wong
Rutgers University
125 Paterson Street
New Brunswick, New Jersey 08901
wongst@rwjms.rutgers.edu

Wanpracha Chaovalitwongse
University of Washington
3900 NE Stevens Way
Seattle, Washington 98195
artchao@uw.edu

ABSTRACT

Epilepsy is one of the most common neurological disorders. The sudden and spontaneous occurrence of epileptic seizures imposes a significant burden on patients with epilepsy. If seizure onset can be predicted effectively, it would greatly improve the life of patients with epilepsy, and also open new therapeutic avenues for the treatment of epilepsy. In contrast to the existing techniques, which mainly make predictions based on some inflexible user-specified thresholds, we propose an adaptive online pattern learning framework to achieve accurate personalized seizure prediction. In particular, we propose a new feature extraction approach to characterize electroencephalogram (EEG) patterns. By monitoring intracranial EEG signals in a sliding window, a pattern library was constructed incrementally online for each patient. A probabilistic prediction framework was developed to make personalized seizure prediction based on the continuously-updated patient-specific EEG pattern libraries. The proposed feature extraction and personalized prediction framework achieved very impressive prediction results on 10 patients with epilepsy. With best parameter settings, the proposed prediction framework achieved very promising with a sensitivity of 89% and a specificity of 87% averaged over the 10 patients. The proposed EEG feature extraction techniques and the probabilistic prediction framework offer a promising diagnostic tool of brainwave signals to solve the challenging seizure prediction problem.

Keywords

EEG Feature Extraction, Brain Abnormality Diagnosis, Cognitive Pattern Recognition, Epileptic Seizure Prediction, Probabilistic Decision Model

1. INTRODUCTION

Epilepsy is one of the most common neurological disorders, affecting approximately 1% of the world's population [6]. Epileptic seizures generally occur without warning, and the shift between a normal brain state and the seizure onset is often considered an unpredictable phenomenon. The unpredictability of seizures represents a significant source of morbidity in patients with epilepsy. These patients frequently suffer from seizure-related injuries due to a loss of motor control, a loss of consciousness or a delayed reactivity during seizures [15]. The ability to predict the occurrence of impending seizures could significantly improve the life quality of patients with epilepsy.

The pioneering efforts to investigate the predictability of seizures were made by Viglione and Walsh in the 1970s [24] and Iasemidis et al. in the 1980aŕs [8]. Since then, many studies of seizure prediction have been carried out due to the great advances in computer processing and the availability of sophisticated mathematical techniques for biological signals. In general, most of the current seizure prediction methods involve two steps. First, univariate or multivariate EEG features are extracted from a sliding window. Then each EEG epoch in the moving window is classified as either pre-seizure or normal based on an optimized threshold level. Whenever a windowed EEG epoch is classified as pre-seizure, a warning alarm is triggered indicating that an impending seizure may occur within a pre-defined prediction horizon. Although some methods have shown good results for selected patients, the reliability and repeatability of the results have been questioned when tested on other EEG datasets. For example, many of the earlier optimistic findings were irreproducible or achieved poor performance in extended EEG datasets [1].

Unlike many biological detection problems with relatively clear patterns to recognize, the pre-seizure EEG patterns are unknown. Given the heterogeneity of the epileptogenic regions of brain and intracranial electrode placement that is individualized per patient, pre-seizure EEG patterns may vary a lot over time for a same patient, and may be dramatically different among patients. Therefore, a significant challenge of seizure prediction is the high inter- and intra-individual variability of epileptic seizures with a variable degree of success [9]. Although many nonadaptive methods have achieved promising results, this high variability makes it difficult to develop a universal robust predictor to accurately predict seizures for a wide range of patients with different seizures. Manually tuning a threshold level for each individual patient is a subjective procedure and would pose a significant

*BrainKDD 2014 New York, USA.

burden on physicians and patients. The inability to apply these techniques to a wide spectrum of epileptic patients with a variety of types of epileptic seizures may represent the greatest limitation of current seizure prediction methods. Therefore, there is an urgent need for an automated adaptive framework for epileptic seizure prediction.

A number of adaptive seizure prediction algorithms have been proposed to account for the high inter- and intra-individual variability of epileptic seizures [9, 10, 21, 19, 4]. Iasemidis et al. [9, 10] and Sackellares et al. [21] developed optimization-based prediction algorithms which, based on dynamical synchronization in the human epileptic brain, adaptively selects a group of critical EEG electrodes to predict impending seizures. Rajdev et al. [19] also proposed an adaptive prediction algorithm based on a Wiener implementation of autoregressive (AR) modeling. A warning was issued if the prediction errors over a moving window exceeded a threshold. The threshold was continuously updated online, and it was optimized to maximize sensitivity and latency, while minimizing FPR. This algorithm achieved an averaged sensitivity of 92% on four rats with 70 seizures. This study also compared the proposed algorithm with the state-of-the-art seizure prediction algorithms [10, 11, 18, 12, 13, 14]. In particular, we are interested to compare the two most recent adaptive algorithms in Rajdev et al. [19] and Iasemidis et al. [7]. It is noted that the FPR in [19] was 4.8/hour, which is much higher than that in [10] (0.18/hour). And the averaged warning time in [19] is only 6.7 seconds, which is much shorter than that in [10] (67.6 minutes).

Given our accumulated knowledge regarding seizure prediction, we conjecture that a promising approach may be the one that processes intelligent learning ability and autonomously adapts to individual patient's EEG patterns. In the past, we have developed online automated seizure prediction algorithms that could adaptively select critical electrodes to predict the next impending seizure onset [9, 2, 21]. However, only adaptively selecting EEG channels limited the prediction performance improvement. In a later study, we also developed a reinforcement learning framework to achieve patient-specific online seizure predictions [26]. This adaptive framework was capable of improving the prediction performance over time as the system learns more about the EEG patterns of a patient. However, it was still restricted to some pre-defined parameters, such as the sample sizes for the normal and pre-seizure baselines. The optimal values for each patient are unknown.

In this study, we have made a great progress to achieve a more efficient personalized online monitoring and seizure prediction framework. We made a number of important contributions including a new time series feature extraction, a new high-level pattern representation, and a new probabilistic pattern prediction scheme. We formulate the online seizure prediction problem as an online adaptive learning framework, which neatly combines the probabilistic theory, the adaptive learning theory and the new feature extraction and representation techniques into an adaptive online seizure prediction system. The proposed new probabilistic online prediction framework can efficiently process massive non-stationary EEG data, and summarize millions of complex time series patterns online at a low computing cost. It has a great potential to realize a personalized accurate seizure prediction for each individual patient.

This paper is organized as follows. In section 2, a newly developed time series feature extraction technique and a high-level pattern representation approach are presented. The probabilistic adaptive online seizure prediction framework, and the evaluation metrics of prediction performance are presented in section 3. The experimental results are provided and discussed in Section 4, and we conclude the paper in Section 5.

2. MATERIALS AND METHODS

2.1 EEG Data Collection and Preprocessing

We used a dataset containing long-term continuous intracranial EEG recordings from ten epileptic patients. The EEG recordings consist of 26 standard channels. Recording durations ranged from 3 to 13 days. Expert epileptologists annotated the EEG recordings to determine the number of seizures, their onset, and their offset points. The characteristics of the ten patients and the EEG data statistics are outlined in Table 1. The placement of the EEG electrodes is shown in Figure 1, which is a modified image of the inferior transverse view of the brain from Potter [16].

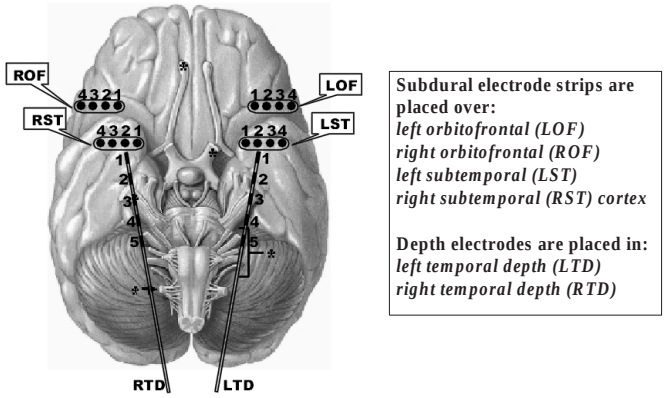


Figure 1: The interior transverse view of the brain and the placement of the 26 EEG electrodes.

Since EEG signals are highly nonstationary and seemingly chaotic, there has been an increasing interest in analyzing EEG signals in the context of chaos theory [20]. Several commonly used chaotic measures in many recent studies include largest Lyapunov exponent [9], correlation dimension [22], Hurst exponent [5] and entropy [17]. Among these EEG measures, the Lyapunov exponent has been shown to be useful in characterizing a chaotic system [23]. In particular, the largest Lyapunov exponent is an important indicator to characterize a chaotic system [23]. In our previous studies, we developed an estimation algorithm called the short-term largest Lyapunov exponent (STL_{max}) to quantify EEG dynamics [9]. We also employ this measure in the current study. A detailed calculation of STL_{max} can be found in Iasemidis [8].

Table 1: The characteristics of the ten patients and the EEG data statistics. (Onset region: LH, left hippacampal; RH, right hippacampal. Seizure types: CP, complex partial; SC, subclinical; GTC, generalized tonic/clonic.)

Patient	Gender	Age (years)	EEG Duration (days)	Number of Seizures	Seizure Rate (per hour)	Seizure Type	Onset Region
1	F	45	3.55	7	0.082	CP, SC	RH
2	M	60	11.98	7	0.024	CP, GTC, SC	RH, RF
3	F	41	8.55	22	0.104	CP	RH
4	M	19	13.13	17	0.054	CP, SC	RH
5	M	33	12.24	18	0.061	CP, SC	RH
6	M	38	3.18	9	0.118	CP, SC	RH
7	M	44	6.09	23	0.157	CP, SC	LH, RH
8	M	29	6.07	19	0.130	CP, SC	RH
9	F	37	11.53	20	0.061	CP, SC	LH, RH
10	M	37	9.88	12	0.051	CP, GTC	LH, RH
total	-	-	86.20	154	-	-	-

2.2 Online Monitoring & Prediction Framework

In this study, we employ an adaptive online learning and prediction framework to discover hidden predictive patterns for epileptic seizures. The flowchart of the online prediction scheme is shown in Figure 2. The proposed framework has the following significant components:

- a two-level feature extraction method. We apply a new feature extraction technique in a sliding-window fashion for online feature extraction from EEG.
- a feature selection technique to select the most important features prior to seizure onset.
- a pattern library to store feature vectors per patient. For each stored pattern, its occurrence frequency and the time related to impending seizures are also stored.
- three adaptive prediction schemes to explore the stored patterns and generate online prediction rules adaptively as the pattern library evolves over time.

These key components of the online monitoring and prediction framework are discussed in detail in the following subsections.

Each STL_{max} point represents the characteristics of a 10-second raw EEG time series (2640 raw EEG points). With the help of STL_{max}, the dimensionality of original time series has been largely reduced. In our previous studies and many other current studies in the world, researchers are trying to perform pattern recognition and prediction based on the converted time series of STL_{max}. However, in practice we notice that online manipulating STL_{max} is still computational expensive. And also it is still extremely complicated to investigate raw STL_{max} time series patterns. This motivates us to develop a new feature extraction technique for time series data, which will be discussed in the next section.

2.3 Two-Level Feature Extraction

As shown in Figure 2, we employ a two-level sliding window approach to extract temporal characteristic features of a multichannel time series epoch. The two-level feature extraction method is presented next.

2.3.1 First-Level Feature Extraction

The first-level characteristic features are extracted directly from each epoch of the raw EEG time series. For each epoch, 29 first-level features are extracted in the first-level sliding window. These are 26 univariate features (the averaged STL_{max} values of the 26 time series channels) and three bivariate features (averaged pairwise Euclidean distance, T-statistic, and Pearson correlation over all channel pairs). We then introduce second-level features to characterize the temporal evolutions of the first-level features.

2.3.2 Second-Level Feature Extraction

Given a time series of a first-level feature, we first applied a piecewise linear approximation algorithm to partition the time series into piecewise linear segments using its key-turning points. We have developed a reliable and efficient algorithm for piecewise linear segmentation of time series data, called two-stage-top-down (TSTD) approach. A more detailed discussion of this algorithm can be found in [25]. After piecewise linear segmentation of a time series $X = (x_1, x_2, ..., x_n)$, its key-turning points become prominent (shown with black dots in Figure 3). There are six linear segments to describe the original time series. Three segments (a, c, e) have increasing trends, and the other three segments (b, d, f) have decreasing trends. Then, the extracted increasing and decreasing trends characterize the temporal fluctuation pattern in a time series. The following four second-level features are proposed to capture the temporal fluctuation of first-level feature time series.

- Feature 1: accumulated vertical decrease in the segmented piecewise linear time series, which is calculated as
$$F_1 = H(a) + H(c) + H(e), \quad (1)$$
where the function $H(.)$ means the vertical distance from the starting point to the ending point of a subsegment in the segmented time series.
- Feature 2: accumulated vertical increase in the segmented piecewise linear time series, which is calculated as
$$F_2 = H(b) + H(d) + H(f), \quad (2)$$
- Feature 3: percentage of the decreasing sub-segments in the time series, which is calculated as
$$F_3 = T(a + c + e)/T(X), \quad (3)$$

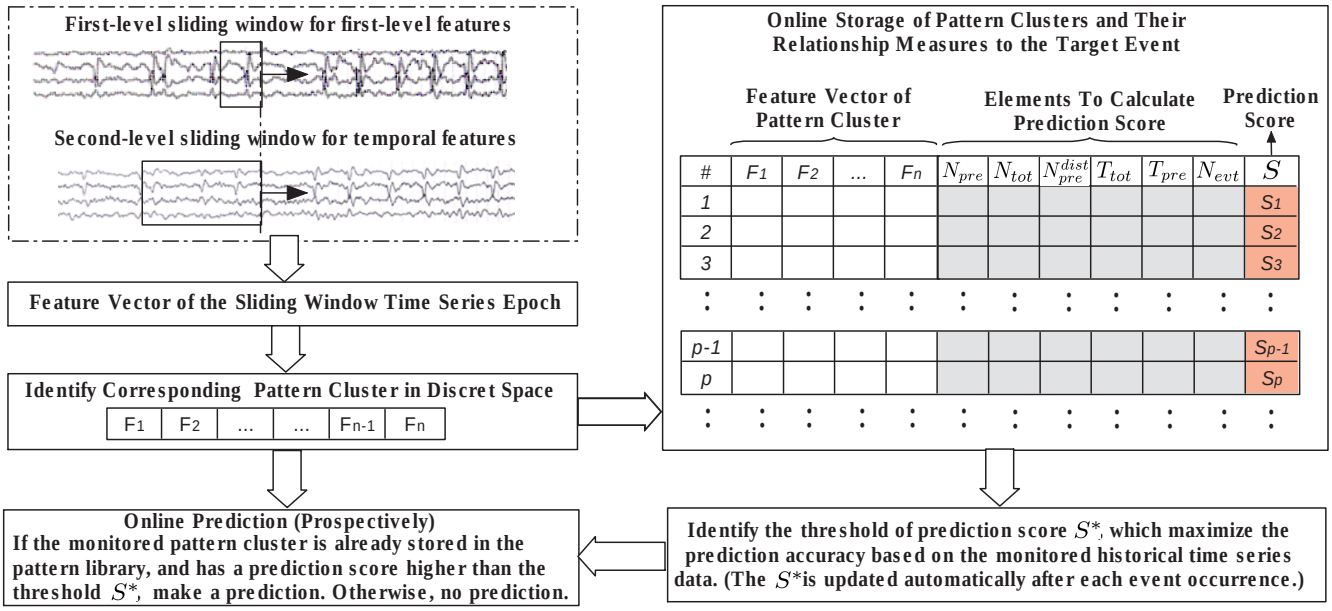


Figure 2: Flowchart of the online learning and prediction framework for personalized seizure prediction.

where $T(.)$ is the horizontal distance from the starting point to the ending point of a sub-segment.

- Feature 4: range of the time series, which is calculated as

$$F_4 = max(X) - min(X), \tag{4}$$

where $max(X)$ and $min(X)$ means the maximum and minimum values of the time series, respectively.

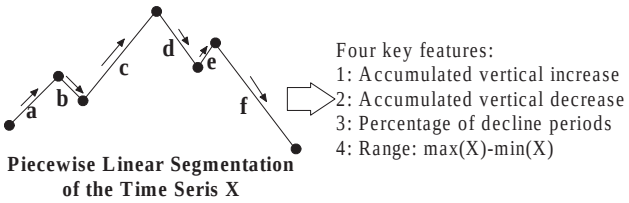


Figure 3: Four key features to represent a time series fluctuation.

With the four features F_1, F_2, F_3, and F_4, we partition each feature space into a number of non-overlap intervals. The time series patterns that fall in the same interval in each feature space represent a set of close-by patterns with similar statistical properties. We consider a set of such time series patterns as a pattern cluster. The concept of pattern cluster is illustrated in Figure 4. The two time series can be represented by the same pattern cluster, namely 1325.

Using the concept of pattern cluster, one can represent millions or billions of time series patterns by a fixed number of pattern clusters representing groups of similar time series patterns. As shown in the example, there are four features, and each feature space is partitioned into five intervals, then the total number of pattern clusters is only $5^4 = 625$ for a single time series. For multivariate time series, one can concatenate the features of each time series into a big feature vector. For example, if there are two time series, the total number of features becomes $4 \times 2 = 8$, and the total number of pattern clusters becomes $5^8 = 625^2$.

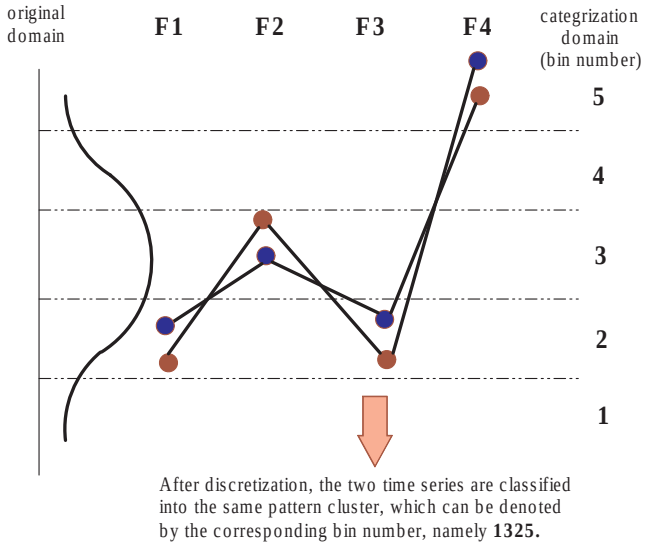


Figure 4: A demonstration of pattern clusters based on discretized feature space.

With this new high-level representation technique, we are capable of dealing with numerous complicated time series patterns by a limited number of pattern clusters. This property is really attractive to analyze chaotic non-stationary time se-

ries patterns. We do not need to worry about an increasing database of recorded pattern clusters, since the maximum number of possible pattern clusters is known fixed number.

2.4 Transformation of Feature Time Series for Online Monitoring

For each EEG epoch monitored online, we have 29 first-level features, and for each first-level feature we have 4 second-level features to describe temporal variations of the first-level features. Then each time series epoch has $29 \times 4 = 106$ features after the second-level feature extraction. To achieve fast learning, it is extremely desirable to monitor only a few 'critical and informative' time series instead of the 26 channels of STL_{max} time series. Fortunately, with extensive research on this problem, we finally extracted two critical time series from the 26 channels of STL_{max} time series as follows:

- The averaged time series of the 26 channels of STL_{max}, called STL_{ave}. The averaged time series could eliminate the random fluctuations among channels and enhance the common or similar patterns exist in most channels. Thus, this transformed time series represents the overall 'common' temporal patterns among the 26 STL_{max} time series.
- The averaged time series of pairwise difference for the 26 channels of STL_{max}, called D_{ave}. We calculate the differences between each pair of STL_{max} time series, and generate a difference time series for each pair. A total number of 325 difference time series are generated for the 26 channels of STL_{max}. The averaged pairwise difference time series represents the overall dissimilarities among the 26 STL_{max} time series.

By using the above time series extraction, we only need to monitor the two transformed time series STL_{ave} and D_{ave}. In a moving window, the 4 second-level features were used to characterize each of the two time series patterns. Thus, each EEG window-epoch were transformed into 8 features and stored in the pattern library. Future, we discretize each feature space into five intervals, then the maximum number of pattern clusters is $5^{(4\times 2)} = 5^8$. The pattern cluster representation has a very low dimensionality, and thus it allows a very efficient storage, visualization, and computational analysis. More importantly, it becomes possible to apply probabilistic theory to analyze the predicability of pattern clusters. In the next, we propose a probabilistic prediction framework to discover the hidden pattern clusters that are predictive to seizure onset.

3. PROBABILISTIC PREDICTION FRAMEWORK

A pattern library with 8-dimension feature space is constructed for each patient. In addition, as shown in the Flowchart in Figure 2, 'pattern information' is stored in the pattern library. This includes pattern's occurrence frequency and occurrence times related to seizure onset. If a pattern appears in pre-seizure periods, it is labeled as a pre-seizure pattern; if it is in normal periods, we can consider it as a normal pattern. The adaptive learning of the predictive power of the monitored pattern clusters is of vital importance in our prediction framework. In this section, we present the probabilistic formula in detail to calculate the predictive score of each pattern cluster. Given a time series pattern cluster, indexed as the kth cluster in the pattern recording table, its prediction score S_k is defined as follows:

$$S_k = \frac{N_{pre}/N_{tot}}{R_{pre}} + \frac{N_{pre}^{dist}}{N_{evt}}, \tag{5}$$

where N_{pre} is the number of occurrences of the pattern cluster in all monitored pre-seizure periods; and N_{pre}^{dist} is the number of pre-seizure periods such that the pattern cluster appears at least once in each of them; N_{tot} is the total number of occurrences of the pattern cluster; and N_{evt} is the total number of seizures that have occurred. For example, if two seizures have been monitored, a pattern cluster occurs three times in the first pre-seizure period, 2 times in the normal periods, and does not show up in the second pre-seizure period, then $N_{pre} = 3$, $N_{pre}^{dist} = 1$, $N_{tot} = 5$, and $N_{evt} = 2$. Finally, R_{pre} is the time ratio between pre-seizure periods and normal periods. In particular, it is calculated as follows:

$$R_{pre} = \frac{T_{pre}}{T_{tot} - T_{pre}} = \frac{N_{evt} \times T_{hrzn}}{T_{tot} - N_{evt} \times T_{hrzn}}, \tag{6}$$

where T_{pre} is the total length of monitored pre-seizure periods, T_{tot} is the total length of monitored EEG time series; and T_{hrzn} is the length of prediction horizon.

The predictive score proposed in formula 5 indicates how strong a pattern cluster is associated with seizure onset. In particular, in the first term of formula 5, the N_{pre}/N_{tot} is the percentage of the pattern cluster appear in pre-seizure periods. This percentage value is compared with R_{pre} to evaluate if the pattern cluster occurs in pre-seizure periods at a random level. If the pattern cluster occurs equal-likely in both pre-seizure and normal periods, then the expected value of N_{pre}/N_{tot} should be equal to the expected value of R_{pre}. In particular, we can summarize the following properties of the first term of formula 5:

- If the pattern is pure random in both pre-seizure periods and normal periods, then
$$E(N_{pre}/N_{tot}) = E(R_{pre}). \tag{7}$$
- If the pattern occurs more frequently in pre-seizure periods than the normal periods, then we have
$$E(N_{pre}/N_{tot}) > E(R_{pre}). \tag{8}$$
The higher the ratio value, the more likely the pattern cluster is associated with seizure onset.
- If the pattern occurs less frequently in pre-seizure periods than the normal periods, then we have
$$E(N_{pre}/N_{tot}) < E(R_{pre}). \tag{9}$$

As discussed above, the ratio of N_{pre}/N_{tot} and R_{pre} (the first term in formula 5) is an important factor to identify the prediction power of a pattern cluster. However, it is noted that this ratio alone is sometime unreliable and un-robust under some extreme situations. For example, a pattern cluster occurs many times within one prediction horizon (may due to noises or unusual situations), and appears much less

frequently or never occurs in other pre-seizure periods. In such cases, the ratio can be temporally high due to its very high occurrence frequency in only a few pre-seizure periods. And thus lead to a high predictive score. Although the ratio may return toward its expected value in long-term if N_{nm} could increase over time. However, it may take a long time and many false predictions may have been made during this period due to this 'bad' pattern cluster.

To remedy this limitation, we introduce the second term in formula 5, N_{pre}^{dist}/N_{evt}, which considers the percentage of the pattern occurrences in different pre-seizure periods. Ideally, we assume that a good candidate for prediction should appear in a large portion of the monitored pre-seizure periods, not only in one or in a few of them. In particular, we expect an ideal predictive pattern cluster should have the following property:

$$\frac{NN_{pre}^{dist}}{N_{evt}} \approx 1, \tag{10}$$

which means that the pattern cluster occurs in almost all of the monitored pre-seizure periods.

The weighted summation of the first and the second term in formula 5 estimates the likelihood of a pattern cluster in pre-seizure periods and reduces the bad effects of some extreme situations. In general, the higher the prediction score, the higher probability the pattern cluster appears in pre-seizure periods, and thus the more prominent it is to predict seizures.

3.1 Online Prediction Procedure

With the probabilistic formulation of EEG patterns, the proposed prediction framework is designed to discover the important pattern clusters that have high probabilities associated with impending seizure onset. The following five key parts define the basic structure of the sliding window based monitoring and prediction framework.

- window size: this is to define the length of time series to be analyzed at each time step.
- moving step length: this is to define a fixed period between two successive moving window.
- prediction horizon: this is equivalent to define the length of the pre-seizure period. At each time stamp, how far we can predict for the next seizure onset.
- the threshold R_{kp}^{2*}: this is to set up the approximation accuracy for the top-down time series decomposition algorithm to extract time series key-turning-points. A lower value of R_{kp}^{2*} leads to fewer key turning points, but coarser approximation for time series temporal patterns.
- the formula to identify the predictive power of a monitored pattern cluster. With more and more monitored pattern clusters, we propose a formula to dynamically calculate the probabilistic relationship between the monitored pattern clusters and seizure onset online.

The working mechanism of the proposed online prediction framework can be summarized into four steps:

- Step 1: Define the online monitoring and prediction settings, including the prediction horizon, the sliding window size, the moving step length, the length of pre-seizure period, the threshold value for the top-down time series decomposition algorithm, and the threshold value for the predictive score.
- Step 2: Online monitoring of EEG recordings. For each EEG epoch in a window, first transform the raw EEG into STL_{max} time series, and then extract the two critical time series STL_{ave} and D_{ave}. Apply the proposed top-down time series decomposition algorithm to STL_{ave} and D_{ave}; and for each them, four characteristic features are calculated using the equations 1, 2, 3, and 4. In this way, the raw EEG epoch is represented by a feature vector of 8 features. Based on the partitioned feature intervals, the EEG epoch is finally categorized into a pattern cluster.
- Step 3: Adaptive learning the predictive power of the monitored pattern clusters using a pattern recording table. The table records all the monitored pattern clusters and their occurrence frequency in pre-seizure and normal periods. A predictive score of each pattern cluster is calculated according to the formula 5. The basic idea of the predictive score is to compare the probability a pattern cluster appears in normal periods with the probability it appears in pre-seizure periods. Intuitively, if a pattern cluster is more likely to appear in pre-seizure periods than in normal periods, it may have a strong association relationship with an impending seizure onset. The higher the predictive score, the stronger such co-existent relationship.
- Step 4: Online decision making of seizure prediction. The critical predictive pattern clusters are those with high predictive scores. The online prediction decisions can be made based on the predefined threshold value S^* for the predictive scores. If the predictive score of the current monitored pattern cluster is higher than the threshold S^*, a warning is triggered to the patient. Otherwise, it continues to read in EEG data. That is, go to Step 2 for the next step of moving window analysis.

3.2 Evaluation of Prediction Performance

Time Block-Based Sensitivity and Specificity: We label the continuous EEG recordings by a series of time blocks. The block length is equal to the length of prediction horizon (H minutes). The pre-seizure periods were defined as pre-seizure time blocks. The normal period between two seizures were divided into a number equal-sized normal time blocks with a length of H min. For each time block, the prediction outcome of that block can be categorized into one of the following four subsets:

- TP_{blk}: there is at least one warning in a pre-seizure time block.
- FN_{blk}: there is no warning within a pre-seizure time block.

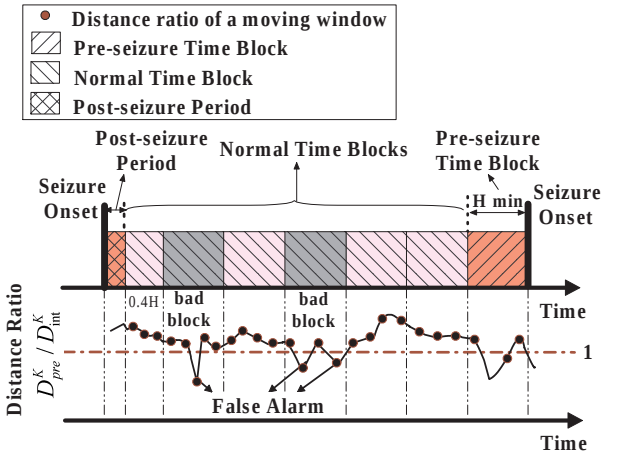
- FP_{blk}: there is at least one warning in a normal time block.
- TN_{blk}: there is no warning within a normal time block.

The time block-based sensitivity and specificity are defined as follows:

$$sen_{blk} = \frac{TP_{blk}}{TP_{blk} + FN_{blk}}, \quad (11)$$

$$spe_{blk} = \frac{TN_{blk}}{FP_{blk} + TN_{blk}}. \quad (12)$$

The time block-based sensitivity and specificity are more suitable to evaluate prediction performance than the traditional definition of sensitivity and specificity, since they consider the effects of prediction horizon for online prediction. A demonstration of the sen_{blk}, spe_{blk} and the traditional sensitivity, specificity is shown in Figure 6.



senblk = 1/1 = 100% (*There is at least one warning within the prediction horizon.*)

speblk= (5.4 - 2) / 5.4 = 62.96% (*The normal period between the two seizures has 5.4 time blocks, two of which are `bad' blocks.*)

traditional sensitivity = 2/4 = 50% (*Percentage of the distance ratios below the threshold 1 within the prediction horizon.*)

traditional specificity = 20/23 = 86.96% (*Percentage of the distance ratios above the threshold 1 in the normal periods.*)

Figure 5: An example to demonstrate the time block-based sensitivity/specificity and traditional sensitivity/specificity.

4. EXPERIMENTAL RESULTS

In this paper, we propose a novel probabilistic framework for a personalized online seizure prediction. In this section, we will evaluate the proposed prediction framework by the EEG recordings from ten patients with epilepsy.

4.1 Prediction Settings

The basic settings of the sliding-window framework are as follows:

- Prediction horizon T_{hrzn} =30, 90, or 150 min, if a prediction is given within T_{hrzn} min prior to a seizure, then we consider that seizure is correctly predicted.
- The scope of the monitoring moving window L_{mw} = 60, 90, or 150 min.
- Step length L_{step} =1, 4, 8, or 16 min (heavily overlapped sliding-window).
- The threshold value of prediction score $S* = 0.4$, 0.7, 1.0, 1.3, 1.6, 1.9, 2.2, 2.5, or 2.8. If a prediction score of a pattern is higher than S^*, then this pattern is classified as a critical predictive pattern.
- Divided the space of each feature into five non-overlap intervals. Since there are four features for a time series, and we have two time series to monitor simultaneously, the total number of possible patterns clusters are $5^4 \times 5^4 = 390625$.

The prediction performance of the framework is evaluated by the block-based sensitivity and specificity, which have been defined in the previous section. Our objective is to find the most important hidden predictive pattern clusters for seizure onset. An adaptive prediction table is employed to record the monitored pattern clusters, and estimate the prediction power of each recorded pattern cluster. The pattern clusters with high scores are used for seizure prediction.

4.2 The Seizure Prediction Performance

From Table 2, one can see that the performance of the proposed online adaptive prediction framework generated very promising prediction results. With the best parameter settings for each patient, the averaged time block based specificity and specificity are both higher than 80%. The results can be summarized as follows:

- The averaged specificity over ten patients is 89.06%, and the averaged false prediction rate (FPR) is as low as 0.136 per hour. The false prediction is very low. This is a big benefit for patients in real-life applications.
- The averaged sensitivity over ten patients is 87.33%, which means that most seizures were corrected predicted. With a high sensitivity, patients are definitely capable of better control their everyday life.

Table 3 summarizes the prediction results for the same 10 patients from a previous adaptive seizure prediction approach proposed by Chaovalitwongse at al. in 2005 [3]. It is noted that the averaged false prediction rates of the two adaptive prediction approaches are at the similar level. However, the sensitivity of our new prediction system is significantly higher than that of the previous approaches. The outcome of this study confirmed that the proposed online monitoring and prediction framework has a superior performance to process massive noisy time series data. An 88% overall prediction accuracy also leads the prediction performance for most of the current existing seizure prediction techniques.

In addition, to demonstrate the fast learning ability of the proposed approach, we also recorded the online sensitivity (OS) over time, which is calculated by

$$OS = \frac{\text{the number of correctly predicted seizures}}{\text{total number of monitored seizure onsets}}. \quad (13)$$

Table 2: The prediction results for 10 patients with epilepsy. The last two columns show the number of discovered critical predictive patterns and the total number of recorded pattern clusters.

Patient	Best Parameter Settings				Prediction Performance			Pattern Information	
	T_{hrzn} (min)	L_{mw} (min)	L_{step} (min)	S^*	sen_{blk}	spe_{blk}	FPR (false/h)	# critical patterns	# all patterns
1	30	120	4	1.3	80.00%	89.38%	0.13	31	387
2	30	60	4	1.9	83.33%	93.22%	0.09	28	748
3	90	60	1	0.4	95.24%	80.77%	0.08	60	598
4	30	120	4	0.4	93.33%	92.23%	0.12	72	582
5	30	60	1	2.8	82.35%	89.87%	0.14	35	854
6	90	120	1	2.8	87.50%	82.14%	0.07	31	219
7	30	120	4	0.4	89.47%	83.94%	0.23	40	500
8	30	90	1	1	93.75%	83.55%	0.21	53	460
9	30	120	1	1.3	94.74%	86.78%	0.16	65	782
10	30	120	1	2.8	90.91%	91.41%	0.13	57	506
	Average Performance				89.06%	87.33%	0.136	47.20	563.60

Table 3: The summary of the prediction results for the same 10 patients from a previous online prediction framework proposed by Chaovalitwongse et al. [3].

Patient	sen_{blk}	FPR (false/hour)
sub1	66.67%	0.1
sub2	66.67%	0.094
sub3	88.89%	0.074
sub4	71.43%	0.19
sub5	25%	0.189
sub6	25%	0.13
sub7	75%	0.181
sub8	71.43%	0.147
sub9	100%	0.178
sub10	80%	0.214
ave.	68.75%	0.154

Take patients 1 as an example, the evolution of the OS values over time is shown in Figure 6. One can observe clearly that there is a strong increasing trend for sensitivity. This indicates that the proposed adaptive framework is capable of discovering complex pre-seizure time series patterns effectively.

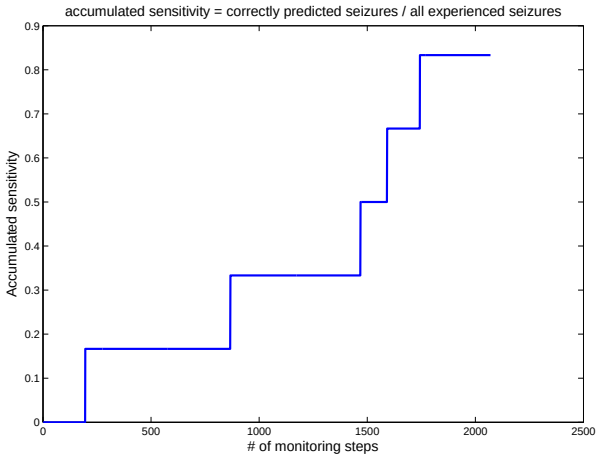


Figure 6: An example to demonstrate the time block-based sensitivity/specificity and traditional sensitivity/specificity.

5. CONCLUSIONS

In this paper, we achieved a very promising seizure prediction performance using the pattern-cluster-based feature extraction techniques and the probabilistic prediction framework. In particular, we made four important contributions in this study. First, we developed a new set of time series pattern features using a robust piecewise-linear segmentation algorithm previously developed by our group. Second, we developed a pattern-cluster based high-level representation to manipulate enormous EEG time series patterns. Third, we proposed a new feature extraction technique, which transforms multichannel raw EEG data into two critical time series for online monitoring. Finally, we formulated the challenging online seizure prediction problem into a novel probabilistic prediction framework using the probability theory. The proposed online monitoring and prediction framework has been applied to perform online seizure predictions for ten patients with epilepsy. The proposed new prediction approach generated impressive superior prediction performance compared with our previous research outcomes and many other seizure prediction studies. The both high sensitivity and high specificity make it possible to achieve a personalized accurate seizure prediction in practical applications. The experimental outcomes of this study are really encouraging considering that seizure prediction techniques are still in their early stages, and many current prediction approaches achieve prediction performances only slightly better than chance.

6. REFERENCES

[1] R. Aschenbrenner-Scheibe, T. Maiwald, M. Winterhalder, H. Voss, J. Timmer, and A. Schulze-Bonhage. How well can epileptic seizures be predicted? An evaluation of a nonlinear method. *Brain*, 126:2616–2626, 2003.

[2] W. Chaovalitwongse, L. Iasemidis, P. Pardalos, P. Carney, D. Shiau, and J. Sackellares. Performance

of a Seizure Warning Algorithm based on the Dynamics of Intracranial EEG. *Epilepsy Research*, 64(3):93–133, 2005.

[3] W. Chaovalitwongse, P. Pardalos, L. Iasemidis, D. Shiau, and J. Sackellares. Dynamical approaches and multi-quadratic integer programming for seizure prediction. *Optimization Methods and Software*, 20(2-3):383–394, 2005.

[4] L. Chisci, A. Mavino, G. Perferi, M. Sciandrone, C. Anile, G. Colicchio, and F. Fuggetta. Real-time epileptic seizure prediction using AR models and support vector machines. *IEEE Transactions on Biomedical Engineering*, 57(5):1124–1132, 2010.

[5] S. Dangel, P. Meier, H. Moser, S. Plibersek, and Y. Shen. Time series analysis of sleep EEG. *Computer assisted Physics*, pages 93–95, 1999.

[6] J. Engel and T. Pedley. *Epilepsy: A Comprehensive Textbook*. Lippincott Williams & Wilkins, Philadelphia, PA, 1997.

[7] L. Good, S. Sabesan, S. Marsh, K. Tsakalis, D. Treiman, and L. Iasemidis. Nonlinear dynamics of seizure prediction in a rodent model of epilepsy. *Nonlinear Dynamics Psychol Life Science*, 14(4):411–434, 2010.

[8] L. Iasemidis. *On the dynamics of the human brain in temporal lobe epilepsy*. PhD thesis, University of Michigan, Ann Arbor, 1991.

[9] L. Iasemidis, D. Shiau, W. Chaovalitwongse, J. Sackellares, P. Pardalos, J. Principe, P. Carney, A. Prasad, B. Veeramani, and K. Tsakalis. Adaptive epileptic seizure prediction system. *IEEE Transactions on Biomedical Engineering*, 50(5):616–627, 2003.

[10] L. Iasemidis, D. Shiau, P. Pardalos, W. Chaovalitwongse, K. Narayanana, A. Prasada, K. Tsakalis, P. Carney, and J. Sackellares. Long-term prospective online real-time seizure prediction. *Clinical Neurophysiology*, 116:532–544, 2005.

[11] K. Lehnertz, R. Andrzejak, J. Arnhold, T. Kreuz, F. Morman, C. Rieke, G. Widman, and C. E. Elger. Nonlinear EEG analysis in epilepsy: its possible use for interictal focus localization, seizure anticipation and prevention. *Journal of Clinical Neurophysiology*, 18:209–222, 2001.

[12] B. Litt, R. Esteller, J. Echauz, M. D'Alessandro, R. Shor, T. Henry, P. Pennell, C. Epstein, R. Bakay, and M. Dichter. Epileptic seizures may begin hours in advance of clinical onset: a report of five patients. *Neuron*, 30:51–64, 2001.

[13] F. Mormann, R. Andrzejak, T. Kreuz, C. Rieke, P. David, C. Elger, and K. Lehnertz. Automated detection of a preseizure state based on a decrease in synchronization in intracranial electroencephalogram recordings from epilepsy patients. *Physical Review E*, 67:021912, 2003.

[14] I. Osorio, M. Frei, and S. Wilkinson. Real-time automated detection and quantitative analysis of seizures and short-term prediction of clinical onset. *Epilepsia*, 39(6):615–627, 2010.

[15] H. Persson, K. Alberts, B. Farahmand, and T. Tomson. Risk of extremity fractures in adult outpatients with epilepsy. *Epilepsia*, 43(7):768–772, 2002.

[16] H. Potter. Anatomy of the brain. http://faculty.ucc.edu/biology-potter/TheBrain/, 2006.

[17] R. Quiroga, J. Arnhold, K. Lehnertz, and P. Grassberger. Kulback-leibler and renormalized entropies: applications to electroencephalograms of epilepsy patients. *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics*, 62:8380–8386, 2000.

[18] M. L. V. Quyen, V. Navarro, M. Baulac, B. Renault, and J. Martinerie. Anticipation of epileptic seizures from standard EEG recordings. *The Lancet*, 361(9361):970–971, 2003.

[19] P. Rajdev, M. Ward, J. Rickus, R. Worth, and P. Irazoqui. Real-time seizure prediction from local field potentials using an adaptive Wiener algorithm. *Computers in biology and medicine*, 40(1):97–108, 2010.

[20] P. Rapp, T. Bashore, J. Martinerie, A. Albano, I. Zimmerman, and A. Mess. Dynamics of brain electrical activity. *Brain Topography*, 2:99–118, 1989.

[21] J. Sackellares, D. Shiau, J. Principe, M. Yang, L. Dance, W. Suharitdamrong, W. Chaovalitwongse, P. Pardalos, and L. Iasemidis. Predictibility analysis for an automated seizure prediction algorithm. *Journal of Clinical Neurophysiology*, 23(6):509–520, 2006.

[22] C. Silva, I. Pimentel, A. Andrade, J. Foreid, and E. Ducla-Soares. Correlation dimension maps of EEG from epileptic absences. *Brain Topography*, 11:201–209, 1999.

[23] J. Vastano and E. Kostelich. Comparison of algorithms for determining Lyapunov exponents from experimental data. In *International conference on dimensions and entropies in chaotic systems*, pages 100–107, Pecos River, NM, USA, 1985.

[24] S. Viglione and G. Walsh. Epileptic seizure prediction. *Electroencephalography and Clinical Neurophysiology*, 39:435–436, 1975.

[25] S. Wang. *Online Monitoring and Prediction of Complex Time Series Events From Nonstationary Time Series Data*. PhD thesis, Rutgers University, 2012.

[26] S. Wang, W. Chaovalitwongse, and S. Wong. A novel reinforcement learning framework for online adaptive seizure prediction. In *Bioinformatics and Biomedicine (BIBM), 2010 IEEE International Conference on*, pages 499–504, dec. 2010.