JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 33, 517-536 (2017)

DAPs: Mining using Change-Point Detection of Epileptic Activity Time Series Data

SUN-HEE KIM[1], LEI LI[2], CHRISTOS FALOUTSOS[3], HYUNG-JEONG YANG[4] AND SEONG-WHAN LEE[1]
[1]*Department of Brain and Cognitive Engineering*
Korea University
Seoul, 136-713 South Korea
[2]*Computer Science Division*
University of California
Berkeley CA, 94720 USA
[3]*School of Computer Science*
Carnegie Mellon University
Pittsburgh PA, 15213 USA
[4]*Department of Computer Science*
Chonnam National University
Gwangju, 500-757 South Korea
E-mail: sunheekim@korea.ac.kr; lilei22@baidu.com; christos@cs.cmu.edu; hjyang@jnu.ac.kr; swlee@image.korea.ac.kr

The goal of this study is to mine meaningful patterns effectively and efficiently via change-point detection of the time series data, with the assistance of domain knowledge and observed data. With those patterns, our method can do segmentation and compression. We developed a novel gray-box approach for mining such data: Domain Assisted Parameter semi-free wave mining (DAPs). DAPs is intended for mining time series with rich domain-specific knowledge based on a chaos model. Specifically, it automatically detects a change-point of time sequences, respecting the minimal description length principle. And the time sequence is segmented based on the detected change-point, and each segment is fitted with a consistent model. The experimental results using both synthetic and real EEG data indicated that the developed method offers a significant improvement in segmentation and compression via pattern detection over other existing methods. DAPs reduced the number of bits of the observed data by detecting the changes in the patterns contained therein and brought about a higher average compression ratio, 1.6% more than WT (level 5). DAPs provides the advantages of (a) being capable of automatically detecting meaningful patterns, (b) being parameter semi-free, and (c) resulting in a huge reduction in data storage. These findings provide possible applications in the use of various medical devices that produce vast amounts of physiological data that should be monitored.

Keywords: segmentation, gray-box model, chaos population model, parameter estimation, minimum description length, electroencephalography, compression

1. INTRODUCTION

The increasing availability of networked medical sensors has allowed for a tremendous amount of physiological data to be continuously gathered, and certain patterns in these data can be used as indicators of patient health. However, the sheer volume of data and the noise contained therein make it impractical to manually monitor patients, who may number

Received August 27, 2015; revised February 9, 2016; accepted October 6, 2016.
Communicated by Vincent S. Tseng.

in the hundreds, each with tens of thousands of continuous signal traces. To this end, data mining techniques have been developed in order to analyze and detect significant patterns in the time series of physiological data and to produce models that represent such data. Change-point detection is a popular technique in the process of mining time series data, the general idea is the identification of abrupt changes in sequential data as an online and offline signal processing step. It is useful in the modelling and prediction of time series in application areas such as disease diagnosis, biometrics, and robotics.

Such a mathematical model can be used to mine the behavior of data from the natural and social sciences, as well as engineering, and it is concerned with simulations of reality reflecting the behavior of dynamic systems [1]. The model tries to mimic the behavior of real systems, be they physical, earth scientific, biological, meteorological, or social systems [2]. A mathematical model of time series data can generally be built from two sources of information: domain knowledge, like physical phenomena or structures; and, observed data from experiments. This allows for two types of modeling: white-box and black-box modeling [3]. White-box models provide exact models by using a-priori knowledge and an assumed structure, and reflect this knowledge in mathematical equations. White-box methods not only require exact domain knowledge but are also highly complex in both implementation and execution.

On the other hand, black-box models can be built by using experimental data and have the advantage of being able to produce simulations very quickly. These models are concerned with deriving a parameter, which optimally fits the observed data. Conventional time-series mining methods often work as black-boxes, which means that they mine patterns or learn models purely from the observational data. Black-box methods are widely used and are readily applicable to a wide range of sequential data. However, black-box models are difficult to interpret, which makes it difficult to achieve phenomenological or qualitative knowledge of a dynamic system [2]. Therefore, in this paper we explore a new perspective by developing a system that mines time series data using gray-boxes.

Gray-box models have some particular benefits relative to both white-box and black-box models [2, 3] because they use both a priori knowledge and observed data. That is, they can reduce parameters by using prior knowledge while still providing a good approximation of the true system. Additionally, such methods assure a low mean-square error because their physical interpretation is possible [4]. As a result, gray-box models are useful and important methods that have many applications in industry, including industrial robots, water environment management, medical science, and sports. In this paper, we propose a novel method, DAPs, which segments and compresses EEG time series data with the assistance of domain knowledge and observed data. Our method improves mining results due to a strategic trade-off between generality and effectiveness. The power of the gray-box approach was examined by studying epileptic behavior in time series data of brain activity.

Epileptic EEG time series data is recorded through multiple sensors in real time, and such data provides information on the electrical activity of brain structures during epileptic seizures [5]. Epileptic EEG time series have piecewise structures with varying characteristics. These piecewise structures may be approximated using linear functions or polynomials [6]. However, linear functions often fail to closely capture patterns in real data because time series data frequently consists of numerous varied patterns [7]. In addition, such functions would require several modeling parameters and be highly sensitive

to the specific choice of parameters. In contrast, the proposed DAPs does not require further user intervention after the initial parameters are automatically set.

In this paper, we demonstrate the effectiveness of DAPs as an expert-built neuron activity model. The empirical results show that DAPs is very suitable at segmenting the epileptic seizure time series data while automatically detecting the patterns. DAPs also shows the capacity of data compression based on its segmentation. The basic idea of DAPs is to detect the change-point of EEG time series data and to split the data into segments that can be separately compressed in order to reduce the number of bits needed to store or transmit data. It employs a gray-box model based on a chaos population model to simulate output signals from input signals [8] and a Minimum Description Length (MDL) method to lead to the best compression rate [9]. The primary contributions of our approach are as follows:

- ***Good Model design*:** DAPs is a novel method that automatically separates changing patterns contained within epilepsy EEG time series data. In effect, it is a parameter semi-free method that can suitably segment data without any additional user intervention after the initial parameters are set.
- ***Effectiveness*:** DAPs can automatically split EEG time series data by using an MDL method. DAPs incrementally finds the best change-points to segment the signals, and the best segmentation is guaranteed to be encoded using fewer bits.
- ***Scalability*:** The run time for DAPs grows linearly with the total input data. Our model can generate a new signal to describe the trends of the input signal. Therefore, DAPs may be adequate for various time series applications where efficient segmentation is needed.

The rest of this paper is organized as follows. Section 2 presents our proposed method to discover patterns by change-point detection. Section 3 shows the experimental evaluation, Section 4 discusses the existing methods in comparison with our method, and Section 5 provides the conclusion.

2. PROPOSED METHOD

In this section we describe the DAPs model that is designed to extract "interesting" patterns from epilepsy EEG time series data, which can do segmentation and compression. To this end, DAPs is proposed as a gray-box model that combines the white-box approach, with which mathematical equations are derived to describe a process in order to perform data analysis, with the black-box approach with which a parameterized model is designed with parameters that are estimated from measurements made on the process itself (with MDL).

2.1 Gray-Box Modeling by the Chaos Population Model

The dynamic neuron population model has been widely used to model neuronal populations and to describe the dynamics of interactions between neurons. Neuronal population models are regarded as effective in describing the dynamic properties of neu-

ronal population activity (see [8] for details). Neuronal population activity is mainly caused by the interactions of groups of neurons. These groups include both excitatory and inhibitory neurons with synaptic connections. Excitatory neurons send inputs to other neurons, which trigger their excitation and bursting. Inhibitory neurons act in the opposite way: they tend to suppress the activity of other neurons. A neuronal population system is composed of two groups of excitatory (e_k) and inhibitory (i_l) neurons. The dynamics of these neurons can be described by means of the dynamical system as follows:

$$\frac{de_k}{dt} = -e_k + \Phi\left(\frac{1}{N_e}\sum_{l=1}^{N_e} a_{kl}e_l - \frac{1}{N_i}\sum_{l=1}^{N_i} b_{kl}i_l - \theta_k^e + p_k\right), \quad k=1, \ldots, N_e, \tag{1}$$

$$\frac{di_k}{dt} = -i_k + \Phi\left(\frac{1}{N_e}\sum_{l=1}^{N_e} d_{kl}e_l - \frac{1}{N_i}\sum_{l=1}^{N_i} g_{kl}i_l - \theta_k^i\right), \quad l=1, \ldots, N_i, \tag{2}$$

where t denotes time, and p_k denotes the external inputs into the excitatory neurons. The parameters a, b, d and g are the strengths of the connections between the populations. The neuronal population models exhibit several types of interactions that involve self- and cross-interactions expressed by a_{kl}, b_{kl}, d_{kl}, and g_{kl}. θ^e and θ^i are the firing thresholds for the excitatory and inhibitory neurons, respectively, and Φ is sigmoidal function.

In this paper, we used the Chaos Population Model (ChaosPM) to simplify dynamic neuron populations with one excitatory and one inhibitory neuronal set. ChaosPM derives its name from the chaotic phenomena exhibited by epilepsy EEG data [10]. The ChaosPM models several of the interactions between excitatory and inhibitory neurons (cross-interactions) and among themselves (self-interactions) as in Fig. 1. ChaosPM uses the average activity of each group as $E_k(t)=1/N_e\sum_k e_k(t)$ and $I_i(t)=1/N_i\sum_l i_l(t)$ from the Eqs. (1) and (2). That is, $E_k(t)$ and $I_i(t)$ represent the activity of the excitatory and inhibitory neurons. This model exhibits oscillations in activity, but it is not difficult to obtain more complex patterns by coupling several of these modules as occurs in the cerebral cortex. These are expressed as C_1, C_2, C_3, and C_4. In addition, the excitatory neurons have an external input, P. C_1 and C_4 are the self-interaction of excitatory and inhibitory neurons, respectively. C_2 is the cross-interaction from inhibitory to excitatory, and C_3 represents the cross-interaction from excitatory to inhibitory. In Fig. 1, solid lines indicate parameter C_1 and C_2 entry on the excitatory neurons and dotted lines denote parameter C_3 and C_4 entering inhibitory neurons.

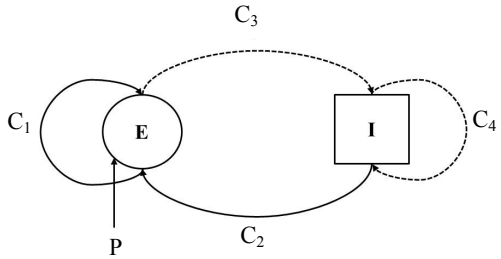


Fig. 1. ChaosPM model: This is coupled by one excitatory (denoted by E) and one inhibitory (denoted by I). The excitatory neuron has the external input.

The activity $E_k(t)$ of an excitatory neuron at location k and time t and the activity $I_k(t)$ of the inhibitory neuron at the same location and time can be described by using Eqs. (3) and (4):

$$\frac{dE_k(t)}{dt} = -E_k(t) + S_\mu(C_1 E_k(t) - C_2 I_k(t) - \Theta^e + P), \tag{3}$$

$$\frac{dI_k(t)}{dt} = -I_k(t) + S_\mu(C_3 E_k(t) - C_4 I_k(t) - \Theta^i), \tag{4}$$

where $E_k(t)$ and $I_k(t)$ denote the activity of the excitatory and the inhibitory neuron at time t, respectively. The thresholds Θ^e and Θ^i should be sufficiently large [11]. S_μ denotes the sigmoid function $S_\mu=[1+\exp^{(-x)}]^{(-1)}$.

ChaosPM uses the above equations to generate a new signal that has t-time points. In detail, we initialize the model with $E_k(0)$ and $I_k(0)$ set to 0. We then obtain new activity for the excitatory and inhibitory neurons, $\hat{E}k(t)$ and $\hat{I}_k(t)$, respectively. The model then creates a new activity signal $\hat{X}(t) = \hat{E}_k(t) - \hat{I}_k(t)$ [12].

The initial C_1, C_2, C_3, and C_4 parameters for ChaosPM are set by using the Levenberg-Marquardt (LM) algorithm [13] to provide initial input parameters. The LM method is a standard technique that can be used to solve non-linear least squares problems. Non-linear least squares methods involve an iterative improvement of the parameter values in order to minimize the sum of the squares of the errors between the function and the observed data points. The parameter values estimated by LM from the sinewave data can be used as the initial input parameters for our proposed method, and the user need not define parameters while running the model for EEG time series data. The preceding learning process for the initial parameter values automatically provides the initial input parameters as follows. The self-interaction of the excitatory neuron C_1 is set to 9.7; the cross-interaction from inhibitory to excitatory neuron, C_2 is set to 10; the cross-interaction from the excitatory to the inhibitory neuron C_3 is set to 5.4; the self-interaction of the inhibitory neuron C_4 is –3.1; and the external input P is 2.5 (see section 3.2 in detail). On the other hand, the Θ^e and Θ^i thresholds are set to constant values 2 and 3.5, respectively. This process occurs as in [11].

2.2 Identification by Minimum Description Length

Our goal is to find patterns contained in epilepsy EEG time series data along with the existing change-points, if any. To this end, two problems had to be addressed: pattern identification and time segmentation. First, given a time series segment $X_{(s)}$ from an epilepsy EEG, good patterns that summarize the fundamental pattern structure of the signals must be found. Second, given an epilepsy EEG time series signal X, the signal segments must be incrementally constructed by selecting the best change-points t_s. Our method is based on the Minimum Description Length (MDL) and can be used to address both of the above problems. MDL is based on the idea that the best model to describe the data is the one that minimizes the sum of the following two items: (1) the amount M of information necessary to describe the model, and (2) the amount of information $D \mid M$ that is necessary to describe the input source using the given model.

$$DL(X)=DL(M)+DL(D|M) \tag{5}$$

where $DL(M)$ is the description length of model M in bits, and $DL(D \mid M)$ is the length of the description of the data when it is encoded with model M in bits. We use a two-part MDL cost for the segmentation and (lossless) compression of the epilepsy EEG time series data. The first part consists of the ChaosPM description costs including those of encoding the parameters that connect each neuron. The second part is the data description cost that consists of the encoding error between the data generated by ChaosPM and the observed data. Our model predicts signals with time t according to a set of given parameters via LM learning, $\theta_p(p=1, \ldots, 5)$. That is, our model (M) requires $log^\star(\theta_p=p+1)$ bits, where $log^\star$ is the universal code (Elias delta) length for integers [9]. However, Elias delta coding cannot code zero or negative integers, so all integers (zeros, positive and negative) are mapped to positive integers as seen in Eq. (6) [14]. The function *PosiInter* normalizes the real-valued signal X into b-bits within a discrete value range. We use 64 values for b so that six bits are sufficient to cover all possible signal values.

$$PosiInter(\hat{X}) = round(\frac{x - \min(x)}{\max(x) - \min(x)}) * 2^b - 1) + 1 \tag{6}$$

Given the length of the signals, t-time, of the epilepsy EEG, we require $\sum_{j=1}^{t} \log^\star(E_j)+1$ to encode the error data. E, given by the model ($D|M$) Therefore, the required number of total bits is as follows:

$$DL(\ddot{X}) = (\sum_{i=1}^{p} \log^\star(\theta_i) + 1) + \sum_{j=1}^{t} \log^\star(E_j + 1). \tag{7}$$

We use the minimum DL to detect patterns in the signal.

2.3 Mining in Time Series

Mining is to discover hidden information or knowledge from the raw data or transformed data, and pattern detection is the most common mining task. In this paper, we detect patterns by identifying the rapidly changing point on continuous time series data that based on MDL. In addition, if the detected patterns have a similarity, then they can be grouped. Otherwise our method separates the different patterns. Given the time series $X=x_1, x_2, \ldots, x_t$, we present an algorithm that constructs signal segments incrementally when a new signal arrives. Intuitively, we want to group "similar" signals from consecutive time series into one signal segment and then encode them all together. For example, signal X^1 and X^2 are similar, and therefore we group them into one signal segment $X_{(s1)}$ while $X_{(s3)}$ is quite different from the previous signal, and hence we start a new segment $X_{(s2)}$. Here, the main principle is still the encoding cost. In particular, our algorithm combines the incoming signal with the current signal segment if there are storage benefits, otherwise we start a new segment with that signal. In this paper, we apply an incremental approach to the significant pattern detection of the epilepsy EEG time series data based on minimum bit cost. Our method estimates the bit cost of the sub-sequence and selects the best time point for pattern identification among those that have the minimum bit cost. To minimize the total number of bits, the time-points for the pattern identification will be decided by using two operations for merging and segmentation. We compute the *Savebit* function that measures how many

bits can be saved by merging and segmenting. *Savebit* is the total number of bits that are saved after applying an operator to create a new slip partition from an existing sub-sequence or by merging with an existing sub-sequence as in Eq. (8).

$$Savebit=DL(Segment)-DL(Merge) \tag{8}$$

Each operator can be defined as follows:

$$Segmenting: (DL(A)+DL(B))\leq DL(C),\ A\neq B, \tag{9}$$
$$Merging: (DL(A)+DL(B))\leq DL(C),\ A\neq B \tag{10}$$

where C is a new sub-sequence created from an existing sub-sequence A and the new sub-sequence B. That is, the number of bits of the existing sub-sequence A, new sub-sequence B, and new sub-sequence C, treated as one sub-sequence from A to B, are computed. We segment these if the sum of each bit of the sub-sequences A and B is bigger than the bits of the new sub-sequence C, hence we start a new segment. On the other hand, they are merged into a signal segment when the new subsequence C takes a larger DL than the sum of the bits for each sub-sequence A and B. DAPs algorithm that is proposed in this paper is expressed in Table 1.

Table 1. DAPs algorithm.

Input: EEG signal, X, and the initial parameters: $PS=(C_1, C_2, C_3, C_4, P)$
Output: Number of segments and total bit costs

for i=1:t ***do*** /* t is the length of a signal
  Create a new signal, $\hat{X}(t)=\hat{E}(t)-\hat{I}(t)$ by Eqs. (3) and (4);
  while minimizing the error cost ***do***
    $ettor$=argmin($\hat{X}$|PS=(C_1, C_2, C_3, C_4, P)/* $error=X-\hat{X}$
    /* Update parameter set
    PS←$PS_{optional}$ /* Optimal parameter set estimated by LM
  end while

  /* minimize description length cost
    Compute $DL(\ddot{X})$ by Eq. (7)
  /* find change-point as extremely high or low picks on the signal
    Compute *Savebit* by Eq. (8)
    if *Savebit* ≥ 0 ***then***
      do Merge
    else
      do Segment
    end if
end for
Return Bit costs, Number of Segmentations, and $PS_{optional}$

3. EXPERIMENTS

3.1 Experimental Data

We evaluated DAPs by using both synthetic and real data. The synthetic data was arbitrarily generated by using ChaosPM with several different initial parameters and without using a learning process for the evaluation. The real data consisted of time series data captured from the measurement of 21 patients with medically intractable epilepsy. The real data were provided by *http://epilepsy-database.eu* (through individual request). This data set was recorded at the Epilepsy Center of the University Hospital of Freiburg, Germany using an invasive recording method. The time series data were obtained from an EEG monitoring system with 6 channels at a sampling rate of 256 Hz. The time series data included 2 to 5 epileptic seizure events over a period of 24 hours (see [15] in detail).

3.2 Parameter Semi-Free Model

Parameter estimation in dynamic models is a central challenge for computational modeling, and it is a difficult and important step in the development of models [16, 17]. In this paper, we used the LM method to solve the parameter estimation problem. This method can automatically detect parameters by an iterative update process of parameters which minimize the error rate between the observed signal and predicted signal. As a result, it is not necessary for users to set parameters whenever the model runs. Merely, the proposed model needs an initial value for parameters before running the process in order to derive the best performance (*i.e.*, it is parameter semi-free). The initial input parameter has a vital influence on the rapidity of convergence. Therefore, we learned to find the best values for initial parameters using LM on the ChaosPM model. It is possible to find initial parameter values by minimizing the mean square error (MSE) between the input data and the predicted data by the model. To decide suitable initial input parameter on ChaosPM, we compared LM and MLE (Maximum Likelihood Estimation). MLE is used the most in parameter estimation methods [18].

Fig. 2 shows the results predicted by ChaosPM when LM and MLE are used to estimate the initial parameters of the model. Figs. 2 (a) and (c) shows the changing accuracy

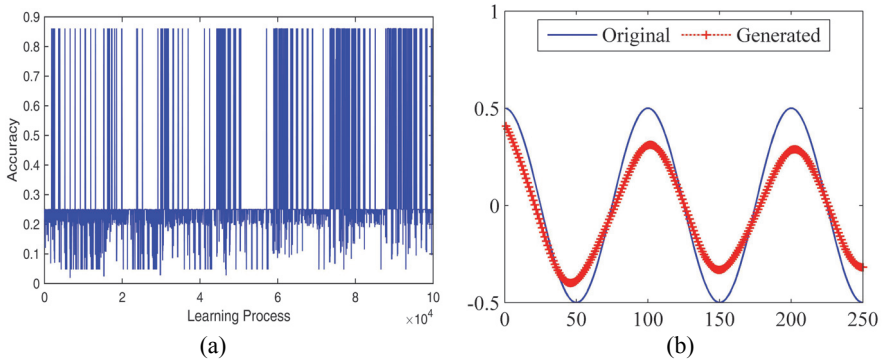
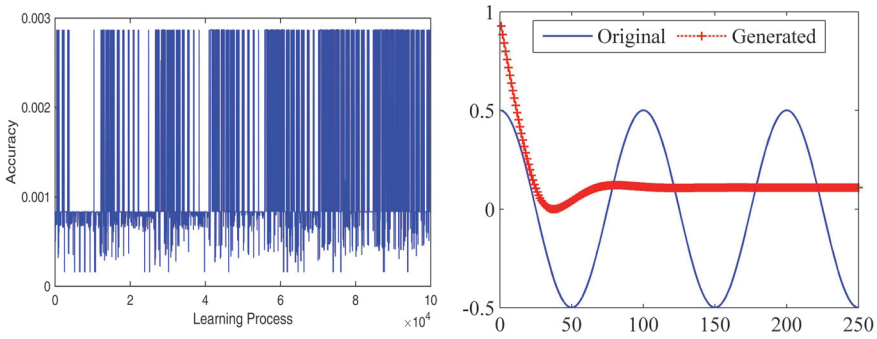


Fig. 2. Comparison between the input signal and the generated signal when using LM and MLE; (a) a learning process by LM; (b) LM result; (c) a learning process by MLE; and (d) MLE result estimated through a learning process of 10^5, respectively.



(c) (d)

Fig. 2. (Cont'd) Comparison between the input signal and the generated signal when using LM and MLE; (a) a learning process by LM; (b) LM result; (c) a learning process by MLE; and (d) MLE result estimated through a learning process of 10^5, respectively.

depending on the learning process of 10^5 using LM and MLE. In (b) and (d) of Fig. 2, the solid line indicates an input sine signal, and the plus-dot line represents the signal predicted by ChaosPM. Fig. 2 (b) shows the results predicted by LM, and Fig. 2(d) shows the result of the MLE. Figs. 2 (b) and (d) are acquired through an iterative process that matches the input signal as closely as possible. Fig. 2 (b) shows a plot of the predicted signal by using the LM with the highest accuracy between the original signal and the predicted signal despite finding only a local minimum, and Fig. 2 (d) shows the signal predicted by the MLE that guarantees the lowest error rates.

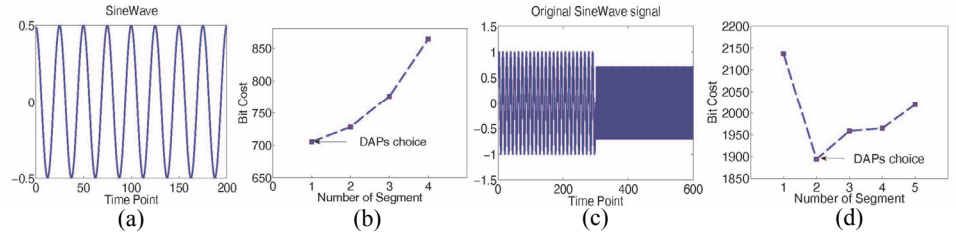
Table 2. Optimal parameters for the initial parameter values.

Method	C_1	C_2	C_3	C_4	P	Accuracy	Loglik/SSE
LM	9.72	9.99	5.42	–3.10	2.53	0.860	4.346
MLE	3.11	6.87	9.09	1.61	3.18	0.002	–602.4

The results confirmed that the LM method provides lower error rates than MLE when the two methods are compared. That is, through the learning process, the LM method found the optimal parameter values with which the predicted signal is most similar to the input signal. Table 2 shows that the optimal parameter values providing high accuracy in the LM and MLE learning process. As a result of Table 2, LM and MLE show high accuracy when C_1, C_2, C_3, C_4 and P parameter values equals those of Table 2. In Table 2, SSE means a sum of squared error between the input signal and the predicted signal by LM, and Loglik shows the maximized log-likelihood value for MLE. These experiments indicate that LM guarantees the best parameter values for the predicted signal. Therefore, the initial parameter values for the DAPs model are automatically set to the best values for the parameters obtained by the LM. This enables DAPs to run without user intervention: it uses the best values that were empirically obtained as the initial parameter values as follows: C_1=9.7, C_2=10, C_3=5.4, C_4=–3.1, and P=2.5. Our proposed method therefore provides a parameter semi-free model that can automatically set the initial parameter values for the DAPs model by using learned parameter values set by the LM.

3.3 Pattern Segmentation Using the Change-Point Detection

In this paper, we performed pattern segmentation via change-point identification based on the MDL with the time series data. Fig. 3 shows the original sine wave signal and the automatically segmented results produced by DAPs. We generated a sine wave signal as seen in Fig. 3 (a), and we used it as input data for the proposed method. The patterns for the signal in order to minimize encoding costs, are incrementally discovered as shown in Table 3, and the length of time needed to compute the minimum bit cost are sequentially increased. Fig. 3 (b) shows how DAPs selects the best number of segments in order to guarantee a minimum bit cost for the signal as a result of the segmenting or merging of the sequences. Therefore, this sine wave signal has no segmentation since it consists of one single pattern. DAPs detected a length of 50 time points as the best change-point from the sine wave signal through the incremental process, and the method performed a merge process for 3 segments from the entire sine wave signal, as shown in Table 3.



(a) (b) (c) (d)

Fig. 3. The DAPs model automatically chooses the best pattern segmentations with the smallest bit cost; (a) Single sine wave signal; (b) bit cost according to the number of segments in a single sine wave signal; (c) mixed sine signal with different frequencies and amplitudes; and (d) bit cost according to the number of segments.

Table 3. Results for the pattern segmenting or merging of sine signals in the DAPs model.

Input Signal	Length of identified pattern	Num. of merge operation	Num. of segment operation	Total length of the signal (time point)
Single Sine Wave (Fig. 3 (a))	50	3	–	200
Mixed Sine Wave (Fig. 3 (c))	60	8	1	600

To verify the pattern segmentation ability of DAPs, we applied a mixed sine wave signal to the model. The mixed signal is composed of two separate sine wave signals with different frequencies and amplitudes, as shown in Fig. 3 (c). Fig. 3 (d) shows the pattern segmentation results obtained through our model. As seen in Fig. 3 (d), DAPs correctly separated the two signals at the point where the frequency and the amplitude changed. In the case of this mixed sine wave signal, the best length to perform pattern identification was detected to have a duration of 60 time points, and was achieved with one segmenting operation and eight merging operations carried out over 10 steps, as shown in Table 3.

In this subsection, we show the pattern segmentation performance measured using a variety of experimental data. We applied DAPs to a synthetic signal created by using Cha-

osPM with several parameter values without the LM learning process. Fig. 4 (a) shows the original synthetic signal combined with the three types of signal. Fig.4 (b) is the result of the segmentation. As shown in Fig. 4 (a), the input signal was created with time points 1 to 300 of the signal using parameter values C_1=3, C_2=7, C_3=10, C_4=–4, and P=3: Time points 301 to 600 were generated using parameter values, C_1=10, C_2=15, C_3=15, C_4=–4, and P=3: The rest signal (time points 601 to 900) was generated by using parameter values C_1=–1, C_2=1, C_3=5, C_4=–5, and P=10. For this synthetic signal, we detected the precise change-points for the pattern segmentation reducing with the number of bits by using MDL. Fig. 4 (b) shows the results of the pattern segmentation incrementally identified by MDL, and Fig. 4 (c) indicates the bit cost according to the number of segments from which DAPs chooses the best number of segments, s(s=3), by segmenting the signal into three pieces. The three types of synthetic signal (see Fig. 4 (a)) require 972 bits in DAPs in order to describe the signal. That is, DAPs provided the smallest number of bits by automatically performing segmenting or merging operations (see Fig. 4 (c)).

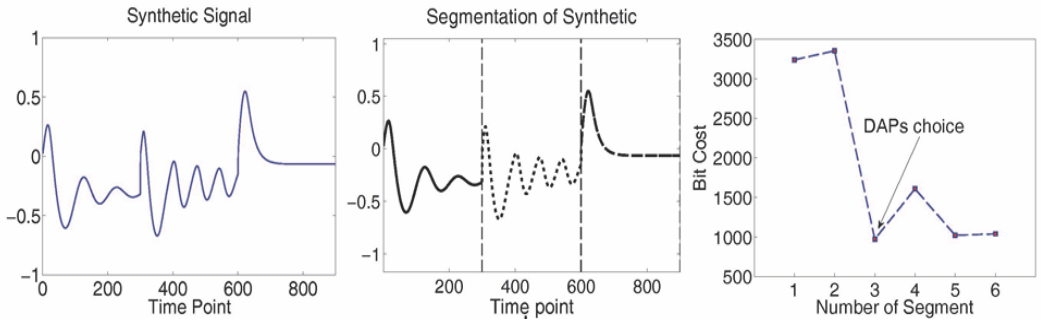


Fig. 4. Pattern segmentation of the synthetic signal consisting of three signal types: (a) the original synthetic signal; (b) the pattern segmented by the DAPs model; and (c) the bit cost according to the number of segments.

Fig. 5 shows the pattern segmentation results for the real signal obtained from patient P4. This patient P4 is 26 age and a female, and seizures of three type as simple partial (SP), complex partial (CP), and generalized tonic-clonic seizure (GTC) appeared over five times during recording EEG signal. EEG signal of P4 was recorded using grid, strip, and depth electrodes [15]. Fig. 5 (a) depicts the original signal of the patient, and includes normal and seizure signals, as defined by expert opinion. We applied DAPs to the real signal, and it detected the time points for a sudden change, as shown in Fig. 5 (b). DAPs then segmented the signal according to the change-points that were detected (changeable pattern). As seen in Fig. 5 (c), DAPs guarantees the minimum bit cost by using ten segments. That is, when EEG signal is split by ten segments, DAPs takes the lowest bit costs near to 20000, and the segmenting and merging processes can identify a reliable pattern to segment a real signal.

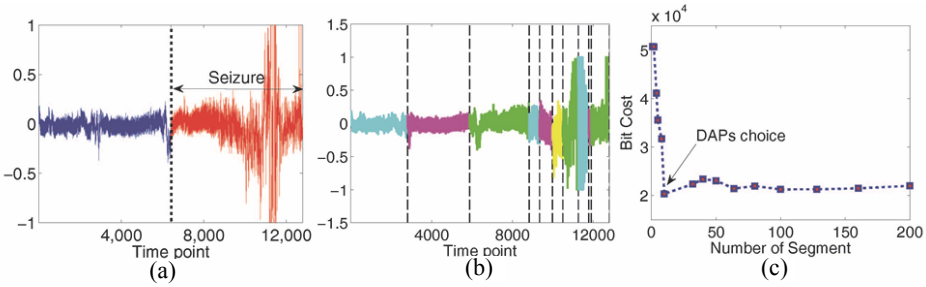


Fig. 5. DAPs automatically chooses s=10 segments for segmenting the real signal; (a) Real EEG signal from patient P4; (b) the result of the pattern segmentation with the proposed model; and (c) the bit cost according to the number of segments.

Fig. 6 shows the results of the automatic segmentation by DAPs of epilepsy EEG data. In Fig. 6, the blue line indicates the original signal, and the multi-colored line indicates the signal reconstructed for the segment separated by DAPs. The vertical dotted line indicates the segment that was automatically separated by DAPs. Fig. 6 (a) is the signal from the first channel of the 5th patient, and Fig. 6 (b) shows the signal of patient P9S1. Patients P5 and P9 are 16 and 44 age, and they are a female and male, respectively. P5 has seizures of three types as SP, CP, and GTC, but P9 only has CP and GTC. During EEG signal recorded, electrodes used grid and strip, and seizures occurred over three and five times, respectively [15]. As shown in Fig. 6, DAPs can also reconstruct the signal that indicates the piecewise structure of the observed signal. As a result, DAPs can detect change-points in order to minimize the number of bits. As a result, our model is able to automatically segment EEG time series data, and can therefore compress data to reduce the space necessary for transmission and storage.

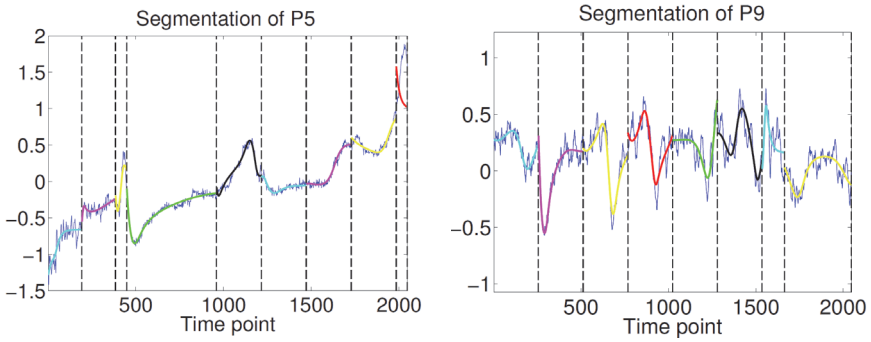


Fig. 6. DAPs pattern segmentation results from real epilepsy EEG data: (a) the signal recorded from the first channel of the patient P5; and (b) the signal recorded from the first channel of the patient P9.

3.4 Effective Compression

Digitized EEG is commonly used for monitoring patients and in patient databases. The volume of these data is necessarily large since a long period of time is required to gather sufficient information from each patient. As a result, large amounts of data have to be either

stored or transmitted making compression necessary to reduce potentially onerous the bit rates [19]. Our proposed method is advantageous that it reduces the number of bits of the observed data by detecting the changes in the patterns contained therein.

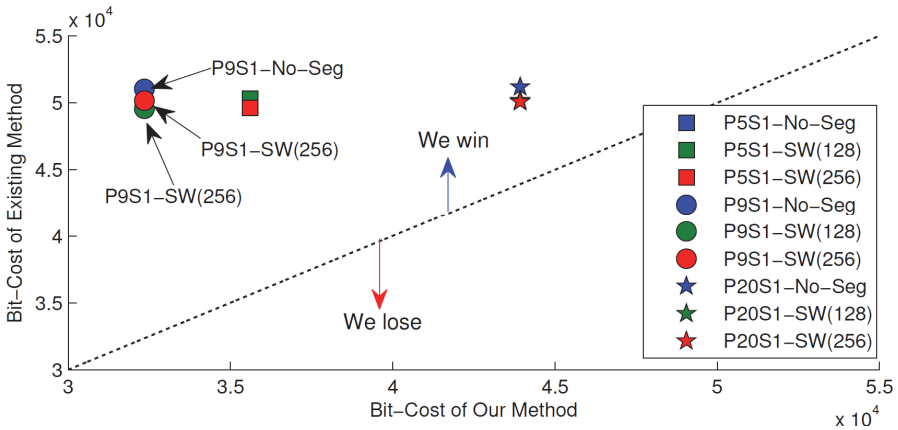


Fig. 7. Bit cost of the total signal via segmentation. We compared the bit cost of the total signals for several patients, including P5, P9, and P20.

Fig. 7 shows the experimental results for the bit cost of several patient signals. We evaluated the performance of the proposed gray-box model by using real EEG data, and showed promising results in terms of the number of reliably identifiable segments while the proposed method compares favorably to a slide window approach [20]. In Fig. 7, the *x*-axis corresponds to the bit cost of our method, and the *y*-axis is the bit cost of existing methods (slide window and no segmentation). In this plot, the black dotted line is considered the standard performance against which the model is evaluated. If the symbols appear along the upper side of the black dotted line, the model provides good performance. On the other hand, if the symbols appear below the line, the proposed method did not perform well. The results indicate that the bit cost of the total of the signals for the data from each patient remain on the upper side of the dotted line. Therefore, DAPs is able to produce a minimum bit cost over the entirety of the signal through segmentation based on the change-point detection. This is considerably lower than the bit cost achieved by the sliding window method and other methods that do not segment the data.

In order to verify the efficiency of the compression, the reduced bit cost obtained by the proposed method was used to measure performance in terms of compression ratio (CR) and the percentage of the root mean square difference (PRD). The CR can be defined as the ratio of the number of bits used to describe the original signal to the number of bits required to represent the compressed signal. PRD is used to evaluate the distortion in the reconstructed signal, and produces the error ratio between the reconstructed and original signals. CR and PRD are together used to evaluate the performance and are calculated as follows:

$$CR = (\textit{Bit of the compressed signal} \,/\, \textit{Bit rate of the original signal}) \times 100(\%), \qquad (9)$$

$$PRD = \sqrt{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2 / \sum_{i=1}^{n} x_i^2} \times 100(\%), \tag{10}$$

where x_i and $\hat{x}_i$ indicate the original signal and the reconstructed signal, respectively. The compression of bio-medical signals is generally performed by using a Discrete Cosine Transform (DCT), Discrete Sine Transform (DST), Fast Fourier Transform (FFT), or Wavelet Transform (WT) [21-24]. In this paper, we have compared the compression ratio performance of our method to that of existing methods, as seen in Table 4. The conventional methods, including DCT, DST, FFT and WT were carried out as MATLAB functions, and the reconstructed signals were obtained by using inverse functions of each method. In the case of the WT, the decomposition was performed at levels 2 and 5 using wavelets db2 and db5.

Table 4. Comparison of the compression ratio between the proposed method and existing methods.

Num. Patient	DCT		FFT		DST		WT				Proposed	
							Level 2		Level 5			
	CR	PRD	CR	PRD	CR	PRD	CR	PRD	CR	PRD	CR	PRD
P1	18.75	0.968	10.46	0.971	18.39	0.960	23.00	0.989	27.48	0.996	37.53	0.953
P2	35.19	0.996	25.02	0.998	35.21	0.998	11.50	0.988	14.60	0.988	39.49	0.986
P3	27.77	0.952	20.89	0.982	27.72	0.957	28.93	1.000	33.08	0.953	47.89	0.960
P4	22.45	0.976	11.84	0.982	22.07	0.965	8.88	0.970	15.30	0.988	48.77	0.967
P5	38.16	0.967	28.36	0.962	38.80	0.981	41.90	0.951	48.06	0.954	39.75	0.997
P6	42.04	0.973	33.11	0.971	42.31	0.989	33.96	0.952	42.18	0.951	37.96	0.953
P7	51.34	0.991	44.73	0.991	50.30	0.962	45.84	0.966	56.46	0.956	32.79	0.979
P8	37.58	0.985	28.91	0.984	37.49	0.988	22.96	0.989	32.94	0.953	40.48	0.963
P9	30.16	0.970	21.56	0.979	30.16	0.965	44.52	0.954	51.42	0.956	27.96	0.988
P10	43.23	0.989	34.22	0.989	43.30	0.985	15.92	0.989	19.22	0.994	39.72	0.999
P11	40.86	0.970	32.89	0.990	41.24	0.980	33.08	0.960	39.66	0.951	48.97	0.993
P12	32.25	0.976	23.38	0.987	31.91	0.953	31.76	0.964	38.40	0.952	41.72	0.991
P13	27.73	0.979	16.45	0.959	27.68	0.956	5.12	0.993	7.74	0.981	43.97	0.952
P14	64.02	0.995	56.55	0.996	63.43	0.982	53.12	0.963	62.06	0.954	28.92	0.953
P15	55.00	0.991	48.39	0.989	52.99	1.000	43.26	0.951	50.44	0.952	29.91	0.981
P16	60.48	0.980	54.20	0.991	60.83	0.998	45.64	0.963	57.30	0.953	49.63	0.986
P17	17.77	0.990	8.40	0.994	17.60	0.971	59.92	0.956	65.70	0.956	50.55	0.970
P18	27.79	0.985	17.39	0.973	27.59	0.963	5.64	0.972	9.42	0.995	42.04	0.973
P19	15.78	0.986	6.45	0.981	15.50	0.978	42.04	0.956	56.04	0.951	28.28	0.958
P20	17.58	0.994	8.02	0.962	17.54	0.975	29.86	0.998	32.94	0.952	16.78	0.964
P21	32.16	0.985	23.33	0.993	32.08	0.983	26.08	0.998	29.16	0.993	49.99	0.953
Average	35.15	0.981	26.41	0.982	34.96	0.976	31.09	0.972	37.60	0.966	**39.19**	**0.972**

Table 4 shows the CR and the PRD for the existing methods and for the proposed method. Generally, the fidelity of the reconstructed signal indicates a distributional range for PRD as follows: 0~2% – 'very good'; 2~9% – 'good'; and 9~19% – 'not good' [25]. In our experiment, we estimated CR by decreasing the PRD values between 0.95 and 1 to compare the compression ratio between the existing methods and DAPs. That is, we compress the signals to obtain a PRD percentage between 0.95 and 1 ($0.95 < PRD < 1$). Therefore, we can interpret the results of the experiments in the following way: if the CR

is high and PRD is low, then a given method may be considered preferable for compressing bio-medical signals. The results in Table 4 indicate that the performance of the existing compression methods achieved compression ratios that were roughly average - 35%, 26%, 34%, 31%, and 37% for DCT, FFT, DST, and WT (level 2 and level 5), respectively, with PRD maintained between 0.95 and 1. These results indicate that our method had a significantly higher average CR, 1.6% more than WT (level 5), compared to these methods. Therefore, our method could have clinical potential for data compression since it achieves better performance than currently available alternatives.

3.5 Scalability for Pattern Segmentation

In this section we experimentally demonstrate the time complexity of DAPs. Fig. 8 shows the running time according to the increase in the total size of several EEG time series. The total size indicates the total time points for a signal, and the complexity of our approach for the segmentation is given as $O(n \cdot r \cdot s)$, where n is the length of the sub-sequence, and r is the number of iterations in the ChaosPM model. Finally, s gives us the number of segments. The results shown in Fig. 8 indicate that the run time grows linearly with respect to the total number of data points.

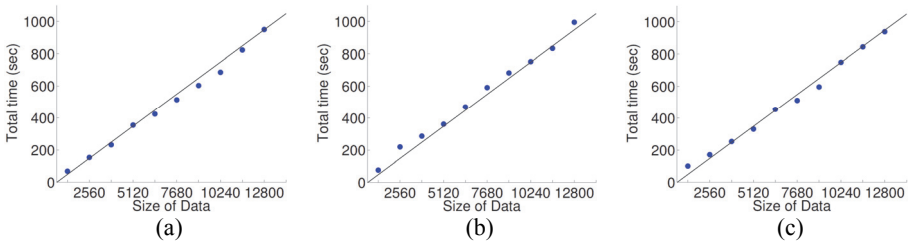


Fig. 8. DAPs: run time of DAPs versus the size of signals in epilepsy EEG data.

4. DISCUSSION

In recent years, there has been a growing interest in using data mining techniques to extract useful patterns from time series data. Time series segmentation is an important issue not only for signal processing but also for data mining. That is, it can be considered as a preprocessing step for variety of data mining tasks or trend analysis techniques. Most time series data that have the characteristic of consisting of continuous numerical values had been dealt with using the fixed-length method for time series segmentation. However, fixed-length segmentation of the time series data may miss significant patterns since meaningful patterns of time series data appear with different lengths. In this paper, we propose a novel method to mine meaningful patterns by detecting the change-point in epilepsy EEG time series data. It can segment time series data by splitting the detected point based on the MDL. Also, in contradistinction to the fixed-segmentation method, the proposed method incrementally separates patterns of different lengths in time series data.

For time series segmentation, Brooks *et al*. [26] introduced a novel segmentation algorithm that can partition quasi-monotonic segments in accordance with scale theory. Piecewise linear representation (PLR) is preferred since it is easier to understand and

implement than other segmentation methods. PLR identifies the change-point that occurs in a time series [27]. However, these methods are not suitable for time series that exhibit sharp fluctuations. Other segmenting approaches for time series data have been proposed, including symbolic mappings, adaptive multivariate splines, hybrid adaptive methods, wavelets, Fourier transforms, and Discrete cosine transform [28-32]. However, none of these can manage time series of different types, nor can they function as parameter-free or semi-free methods. Our method is a parameter semi-free method without user intervention after setting initial parameter values.

Time series data mostly increase linearly with time as patterns are discovered, then there will be a storage problem. Therefore, an effective mechanism for compressing the huge amount of time series data is needed. Pratt and Fink [33] present a compression technique that selects some of the minima and maxima in a series and drops the other points. Recently, Xu *et al*. [34] adopted a sliding window method to compress time series data. This model is based on an MDL/MML method that is in turn based on a fixed window size, and the model captures the information distribution within the data. In this paper, we proposed a novel method for performing incremental pattern detection of epilepsy EEG data with change-points that are identified in a completely automated manner. This method can do compressions of time series data that can help to solve the storage problem.

The performance of the proposed method is verified by comparing the number of bits with the sliding window method that was adopted in [34-37]. We also compared the compression rate against that of existing methods. As seen in section 3, the proposed method provided remarkable results for segmentation by identifying patterns separated by the lowest number of bits that could properly represent the observed data.

5. CONCLUSIONS

We focused on mining patterns via change-point detection of epilepsy EEG time series data in order to predict the epilepsy seizure that is our final goal. To this end, we proposed DAPs, a parameter semi-free method to mine epilepsy EEG time series data. Our method has the following desirable properties. (1) It is rigorous and automatic and can be implemented without additional user intervention (parameter semi-free). (2) It incrementally discovers patterns in time series data by using MDL. (3) It helps reduce the number of bits necessary to completely represent the real data. In addition, the DAPs algorithm grows linearly with respect to the total data size, and therefore, is an adequate and advantageous tool for use in various applications with time series data in order to efficiently process vast amounts of data. In the future work, we will try to prove that the proposed method is able to apply in various fields such as financial data, epidemics data, other bio-signals, and *etc*.

ACKNOWLEDGMENTS

This work was supported by a Korea University Grant. This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (NRF-2015R1D1A1A01057440), and this work was supported by the Brain Korea 21 PLUS Program through the National

Research Foundation of Korea, funded by the Ministry of Education. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2014R1A1A2054530). Also, This work was supported by the Human Resource Training Program for Regional Innovation and Creativity through the Ministry of Education and National Research Foundation of Korea(NRF-2014H1C1A1066771)

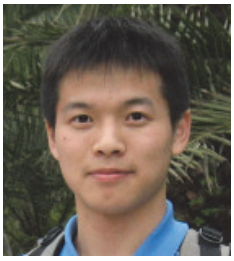
REFERENCES

1. G. Ouyang, C. Dang, D. A. Richards, and X. Li, "Ordinal pattern based similarity analysis for EEG recordings," *Clinical Neurophysiology*, Vol. 121, 2010, pp. 694-703.
2. G. Verghese, "Getting to the gray box: Some challenges for model reduction," in *Proceedings of the American Control Conference*, 2009, pp. 5-6.
3. J. Hauth, "Grey-box modelling for nonlinear systems," Dissertation of Technische University of Kaiserslautern, https://kluedo.ub.uni-kl.de/files/2045/diss.pdf. 2008.
4. L. Ljung, "Perspectives on system identification," *Annual Reviews in Control*, Vol. 34, 2010, pp. 1-12.
5. A. Garcés, L. Orosco, P. Diez, and E. Laciar, "Automatic detection of epileptic seizures in long-term EEG records," *Computers in Biology and Medicine*, Vol. 57, 2014, pp. 66-73.
6. H. Aurlien, I. O. Gjerde, G. E. Eide, J. C. Brogger, and N. E. Gilhus, "Characteristics of generalised epileptiform activity," *Clinical Neurophysiology*, Vol. 120, 2009, pp. 3-10.
7. A. Sierra-Marcos, M. L. Scheuer, and A. O. Rossetti, "Seizure detection with automated EEG analysis: A validation study focusing on periodic patterns," *Clinical Neurophysiology*, Vol. 126, 2015, pp. 456-462.
8. S. H. Kim, C. Faloutsos, and H. J. Yang, "EEG-MINE: Mining and understanding epilepsy data," *Trends and Applications in Knowledge Discovery and Data Mining Lecture Notes in Computer Science*, Vol. 7867, 2013, pp. 155-167.
9. T. C. M. Lee, "An introduction to coding theory and the two-part minimum description length principle," *International Statistical Review*, Vol. 69, 2001, pp. 169-183.
10. K. Ogo and M. Nakagawa, "Chaos and fractal properties in eeg data," *Electronics and Communications in Japan*, Vol. 3, 2007, pp. 27-36.
11. H. G. Schuster and P. Wagner, "A model for neuronal oscillations in the visual cortex 1. Mean-field theory and derivation of the phase equations," *Biological Cybernetics*, Vol. 64, 1990, pp. 77-82.
12. H. R. Wilson, "Spikes decisions and actions: Dynamical foundations of neuroscience," Oxford University Press, NY http://www.math.pitt.edu/~bard/classes/compneuro/Chapter7.pdf, 1999, pp. 149-190.
13. H. P. Gavin, "The Levenberg-marquardt method for nonlinear least squares curve-fitting problems," Department of Civil and Environmental Engineering, Duke University, http://people.duke.edu/~hpgavin/ce281/lm.pdf, 2013.
14. T. Rakthanmanon, E. J. Keogh, S. Lonardi, and S. Evans, "MDL-based time series clustering," *Knowledge and Information Systems*, Vol. 33, 2012, pp. 371-399.
15. B. Schelter, M. Winterhalder, T. Maiwald, A. Brandt, A. Schad, J. Timmer, and A.

Schulze-Bonhage, "Do false predictions of seizures depend on the state of vigilance? A report from two seizure-prediction methods and proposed remedies," *Epilepsia*, Vol. 47, 2006, pp. 2058-2070.
16. J. Tabak, C. R. Murphey, and L. E. Moore, "Parameter estimation methods for single neuron models," *Journal of Computational Neuroscience*, Vol. 9, 2000, pp. 215-236.
17. A. A. Poytona, M. S. Varziri, K. B. McAuleya, P. J. McLellana, J. Ramsay, S. Lonardi, and S. Evans, "Parameter estimation in continuous-time dynamic models using principal differential analysis," *Computers and Chemical Engineering*, Vol. 30, 2006, pp. 698-708.
18. I. J. Myung, "Tutorial in Maximum Likelihood Estimation," *Journal of Mathematical Psychology*, Vol. 47, 2000, pp. 90-100.
19. G. Higgins, B. McGinley, S. Faul, R. P. McEvoy, M. Glavin, W. P. Marnane, and E. Jones, "The effects of lossy compression on diagnostically relevant seizure information in eeg signals," *IEEE Journal of Biomedical and Health Informatics*, Vol. 17, 2013, pp. 121-127.
20. E. Fuchs, T. Gruber, J. Nitschke, and B. Sick, "Online segmentation of time series based on polynomial least-squares approximations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 32, 2010, pp. 2232-2245.
21. K. Ranjeet, A. Kumar, and R. K. Pandey, "Ecg signal compression using different techniques," in *Processings of the Advances in Computing, Communication and Control*, 2011, pp. 231-241.
22. H. Garry, B. McGinley, E. Jones, and M. Glavin, "An evaluation of the effects of wavelet coefficient quantisation in transform based EEG compression," *Computers in Biology and Medicine*, Vol. 43, 2013, pp. 661-669.
23. P. Y. Chen, E. Chiliao, and C. Weiliang, "The segmented-matrix algorithm for haar discrete wavelet transform," *Journal of Information Science and Engineering*, Vol. 24, 2008, pp. 1273-1282.
24. C. H. Chen, V. S. Tseng, H. H. Yu, and T. P. Hong, "Time series pattern discovery by a PIP-based evolutionary approach," *Soft Computing*, Vol. 17, 2013, pp 1699-1710.
25. Y. Zigel, A. Cohen, and A. Katz, "The weighted diagnostic distortion measure for ECG signal compression," *IEEE Transactions on Biomedical Engineering*, Vol. 47, 2000, pp. 1424-1430.
26. M. Brooks, Y. Yan, and D. Lemire, "Scale-based mono-tonicity analysis in qualitative modelling with at segments," in *Processings of the International Joint Conferences on Artificial Intelligence*, 2005, pp. 400-405.
27. Y. Zhu, D. Wu, and S. Li, "A piecewise linear representation method of time series based on feature points," *Knowledge-Based Intelligent Information and Engineering Systems Lecture Notes in Computer Science*, Vol. 4693, 2007, pp. 1066-1072.
28. X. Liu, Z. Lin, and H. Wang, "Novel online methods for time series segmentation," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 20, 2008, pp. 1616-1626.
29. L. A. Tang, X. Yu, S. K. Kim, J. W. Han, C. C. Hung, and W. C. Peng, "Tru-alarm: trustworthiness analysis of sensor networks in cyber-physical systems," in *Processing of IEEE International Conference on Data Mining*, 2010, pp. 1079-210.
30. K. Teymourian and A. Paschke, "Knowledge-based processing of complex stock market events," in *Processings of International Conference on Extending Database Technology*, 2012, pp. 594-597.

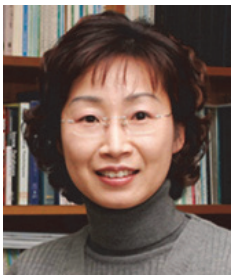
31. K. Srinivasan, J. Dauwels, and M. R. Reddy, "Multichannel EEG compression: wavelet-based image and volumetric coding approach," *IEEE Journal of Biomedical and Health Informatics*, Vol. 17, 2013, pp. 113-120.
32. J. L. Hsu, Y. S. Wu, and I. C. Wu, "A hybrid method for sequence clustering," *Journal of Information Science and Engineering*, Vol. 30, 2014, pp. 1483-1503.
33. K. B. Pratt and E. Fink, "Search for patterns in compressed time series," *International Journal of Image and Graphics*, Vol. 2, 2002, pp. 89-106.
34. K. Xu, Y. Jiang, M. Tang, C. Yuan, and C. Tang, "PRESEE: An MDL/MML algorithm to time-series stream segmenting," *The Scientific World Journal*, Vol. 2013, 2013, pp. 1-11.
35. D. Preston, P. Protopapas, and C. Brodley, "Event discovery in time series," in *Processings of the SIAM International Conference on Data Mining*, 2009, pp. 61-72.
36. C. S. Perng, H. Wang, S. R. Zhang, and D. S. Parker, "Landmarks: a new model for similarity-based pattern querying in time series databases," in *Processing of the 16th International Conference on Data Engineering*, 2000, pp. 33-42.
37. C. H. Chen, T. P. Hong, and V. S. Tseng, "Fuzzy data mining for time-series data," *Applied Soft Computing*, Vol. 12, 2012, pp. 536-542.

Sun-Hee Kim received the B.S. in Multimedia from Korean Educational Development Institute in 2004 and the M.S. degree in Computer Science from Dongguk University, Korea in 2006. She received the Ph.D. degrees in Computer Science from Chonnam National in 2011. She is currently a Research Professor at Department of Brain and Cognitive Engineering, Korea University, Seoul, Korea. Her research interests include data mining, machine learning and bioinformatic.



Lei Li is a Post-Doctoral researcher at EECS Department of UC Berkeley. His research interest lies in the intersection of machine learning, statistical inference and database systems. He received his B.S. in Computer Science and Engineering from Shanghai Jiao Tong University in 2006 (ACM honored class) and Ph.D. in Computer Science from Carnegie Mellon University in 2011, respectively. His dissertation work on fast algorithms for mining co-evolving time series was awarded ACM KDD best dissertation (runner up).

Christos Faloutsos is a Professor at Carnegie Mellon University. He has received the Presidential Young Investigator Award by the National Science Foundation (1989), the Research Contributions Award in ICDM 2006, the SIGKDD Innovations Award (2010), 20 "best paper" awards (including two "test of time" awards), and four teaching awards. Five of his advisees have attracted KDD or SCS dissertation awards. He is an ACM Fellow, he has served as a member of the executive committee of SIGKDD; he has published over 300 refereed articles, 17 book chapters and two monographs. He holds nine patents and he has given over 40 tutorials and over 20 invited distinguished lectures. His research interests include large-scale data mining, for graphs and streams; networks, fractals, and multimedia databases.



Hyung-Jeong Yang received her B.S., M.S. and Ph.D. from Chonbuk National University, Korea. She is currently an Associate Professor at Department of Electronics and Computer Engineering, Chonnam National University, Gwangju, Korea. Her main research interests include multimedia data mining, pattern recognition, artificial intelligence, e-learning, and e-design.

Seong-Whan Lee is the Hyundai-Kia Motor Chair Professor at Korea University, where he is the head of the Department of Brain and Cognitive Engineering and the Director of the Institute for Brain and Cognitive Engineering. He received the B.S. degree in Computer Science and Statistics from Seoul National University, Seoul, Korea, in 1984, and the M.S. and Ph.D. degrees in Computer Science from Korea Advanced Institute of Science and Technology in 1986 and 1989, respectively. His research interests include pattern recognition, computer vision, and brain engineering. He has more than 250 publications in international journals and conference proceedings, and authored 10 books.