Real-time Epileptic Seizure Detection from EEG Signals via Random Subspace Ensemble Learning

Mohammad-Parsa Hosseini, Abolfazl Hajisami, and Dario Pompili
Departement of Electrical and Computer Engineering, Rutgers University–New Brunswick, NJ, USA
E-mails: {parsa, hajisamik, pompili}@cac.rutgers.edu

***Abstract*—Real-time detection of seizure activities in epileptic patients is crucial and can help improve patients' quality of life. Accurate evaluation, pre-surgery assessments, seizure prevention, and emergency alerts for medical aid all depend on the rapid detection of the onset of seizures. A new method of feature selection and classification for rapid and precise epileptic seizure detection is discussed. In this solution, informative components of Electroencephalogram (EEG) data are extracted and an automatic method is presented using Infinite Independent Component Analysis (I-ICA) to select efficiently independent features. The feature space is divided into subspaces via random selection, and multi-channel Support Vector Machines (SVMs) are used to classify the subspaces; then, the result of each classifier is combined by majority voting to find the final output. To evaluate the solution, a benchmark clinical intracranial EEG (iEEG) of eight patients with temporal and extratemporal lobe epilepsy has been considered in a multi-tier cloud-computing architecture. Via the leave-one-out cross-validation, accuracy, sensitivity, specificity, and false positive and false negative ratios of the proposed method are 0.95, 0.96, 0.94, 0.06, and 0.04, respectively, which confirm the effectiveness of the proposed solution.**

***Index Terms*—Brain-Computer Interface; Cloud Computing; Electroencephalogram; Epileptic Seizures; Pervasive Computing.**

I. Introduction

Motivation: Two or more epileptic seizures accruing in a 24-hour period is considered as epilepsy. The prevalence of epilepsy in developed countries has been reported as ranging from 4 to 10 cases per 1000, while that observed in developing countries has been reported as 14 to 57 per 1000 [1]. The best strategy to prevent epileptic seizures is to detect them accurately at onset and to attempt promptly the appropriate therapy (e.g., by administering anticonvulsant drugs or applying brain neurostimulation), which depends on several factors including the frequency and severity of the seizures as well as the person's age, overall health, and medical history. Unfortunately, there are several problems with detecting the onset of epileptic seizures. First, seizure recognition cannot be performed by the patients themselves as many of them do not receive any warning sign before seizures or are not able to press an alarm button during seizures [2]. Second, while neurologists and trained physicians can detect the onset of epileptic seizures by visually scanning large quantities of continuous Electroencephalograms (EEGs), the procedure for doing so is quite complex and time consuming, and it sometimes leads to disagreements among physicians [3]. Third, even when physicians are capable of detecting the onset, they are most likely not available to conduct the diagnosis and administer the therapy quickly enough [4].

In some applications such as responsive neurostimulation [5] a successful therapy depends on the rapid detection of the onset of seizures. Consequently, developing autonomic-computing methods via Brain-Computer Interface (BCI) for epileptic seizure detection would greatly assist clinical care.

Vision: The goal of autonomic computing is freeing the human mind from low-level details to create computing systems with self-management capabilities. By using autonomic-computing concepts in epilepsy health care, we frame the problem as one of designing an autonomic system whose goal is to regulate automatically the human brain so as to prevent it from suffering a full-blown epileptic seizure. Such a closed-loop automatic system can be implemented in two separate steps: the first would consist in developing a BCI to detect onset of epileptic seizures; the second would concern the application of an appropriate neurostimulation signal to remove the seizures. In this study, we focus on the first step, i.e., seizure detection, where a high detection accuracy—especially in terms of false positives and false negatives—is key. For example, a false positive in which the system detects a seizure when there is none may lead to exposing the patient's brain to an unnecessary level of radiation, whereas a false negative would cause the system to fail in detecting and, consequently, preventing an epileptic seizure.

Challenges: For automatic seizure detection, we need to extract appropriate features from the EEG signal and then classify it as normal or epileptic based upon them. Since the scalp has low conductivity, the EEG signal is attenuated and distorted with noise and artifacts, which makes feature extraction and classification prone to errors. To obtain cleaner data, we can record the EEG signal directly from the exposed brain surface, a procedure called "intracranial EEG" (iEEG). Recently, implanted electrodes have been used for recording iEEG in epileptic seizure detection [6]. This method of brain signal recording has some potential advantages such as high spatio-temporal resolution and electro-optic mapping of the dynamic neuronal activity. However, implanted electrodes generate massive amounts of real-time data leading to the big (medical) data problem, which calls for a safe storage to save the large volume of data and for high computational resources to process it at velocity. Moreover, iEEG records a variety of patterns with fluctuations in amplitude and frequency that make feature extraction and classification even more chal-

978-1-5090-1654-9/16 $31.00 © 2016 IEEE
DOI 10.1109/ICAC.2016.57

IEEE computer society

lenging. Although several methods have been developed for detection of epileptic seizures, automatic real-time detection of the ictal (seizure) phase still remains problematic.

Our Approach: To develop an accurate seizure-detection system that is useful in real-life support, we need to perform real-time signal processing, machine-learning computing, and brain-state predictions not only on an individual data set but also on large data sets collected from vast user populations over extended time periods. Therefore, next-generation BCI-EEG systems may be connected to High-Performance Computing (HPC) servers through the Internet to adapt the prediction models to the incoming streaming (in-transit) data.

As a first step towards a system that we believe will ultimately reduce the suffering caused by epilepsy, we propose a cloud-computing framework that automatically detects epileptic seizures. Cloud computing provides a simple way to access storage, databases, and computational resources over the Internet. This growing area of IT services offers ubiquitous access with the potential to increase agility with lower costs. The proposed service platform makes decision based on comparing the extracted EEG patterns with the cloud data. Our seizure-detection framework, as shown in Fig. 1, is composed of: (i) resource provider, (ii) data provider, (iii) service requester, (iv) arbitrator, and (v) cloud storage. Resource and data providers provide computational resources and training data, respectively, for seizure detection. A service requester has the duty of arranging requests for the framework. An arbitrator processes the request from a service requester, determines the set of service providers, and distributes the workload tasks among the resource providers. Finally, cloud storage saves the processed data as history of the patient. Pervasive-computing systems have the great potential of effectively understanding brain activities and can be developed to improve health care in different ways. Using new high data-rate telecommunication technologies and cloud-based pervasive computing, the Quality of Life (QoL) of epileptic patients can be improved tremendously via real-time seizure detection.

Our Contributions: We introduce a cloud-based pervasive data-collection and analysis framework for automatic and real-time seizure detection. Since one of the main goals in seizure detection is extracting the rhythmic nature of brain signals, a wavelet transform is used to divide the signal into different frequency bands. We also extract multiple features from different domains used for classification purposes. Then, we propose a new technique for feature selection based on Infinite Independent Component Analysis (I-ICA). Lastly, we present a new classification method based on ensemble learning and randomness that aims at increasing the sensitivity and decreasing the false detection rate. To sum up, our contributions are:

- A seizure-detection system implemented in the cloud;
- A feature-selection technique using I-ICA, which extracts independent features and infers the number of features automatically from data;
- A random subspace ensemble method using Support Vector Machines (SVMs) as the base classifiers for parallel classification fitting big data problems.

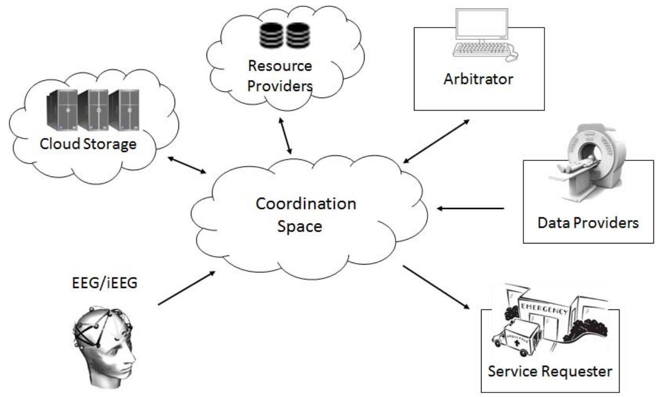

Fig. 1. Cloud-computing framework enabling pervasive healthcare for epileptic patients, which is composed of: resource provider, data provider, service requester, arbitrator, and cloud storage. To adapt the prediction models to the real-time incoming data, next generation BCI-EEG systems may be connected to High-Performance Computing (HPC) servers through the Internet.

To study the classification performance, the system is evaluated and compared against other relevant methods on an available benchmark seizure dataset. The results corroborate the effectiveness of our seizure-detection system as well as its usefulness in real-life support for people with epilepsy.

Paper Outline: The rest of the paper is organized as follows. In Sect. II, the state of the art is presented. In Sect. III, we introduce the system model and detail the architecture, pre-processing, processing (feature selection), and post-processing (classification) steps. In Sect. IV, we validate our assumptions through simulations and show the benefits of our solution over existing ones. In Sect. V, the real-time framework and its application are discussed. Finally, in Sect. V, we draw the conclusions and wrap up the paper by discussing future work.

II. State of the Art

Many algorithms have been recently proposed and modified for seizure detection using EEG and iEEG. Frequency is a key characteristic that has been used in the literature to define abnormality in brain signals. In [7], the authors present a wavelet-based algorithm to examine how different frequency ranges in iEEG fluctuate from the background. Authors in [8] propose an algorithm based on artificial neural networks for classification of EEG signals into healthy, ictal, and interictal. In [9], the authors present an adaptive thresholding technique based on Short Time Fourier Transform (STFT) to find the power spectrum in EEG segments and seizure identification. Using wavelet, in [10] a real-time automated and patient-independent algorithm for detecting absence seizures in WAG/Rij rats as a valid animal model of human absence epilepsy is discussed.

In general, existing works have been developed for local processing and storage without considering multiple channels and big patient data. Since accurate detection of seizures requires analyzing long-term multiple channels, existing works fall short as they do not exploit in full the data available. For a rapid real-time monitoring system, the massive amount of data from different EEG electrodes needs to be stored and

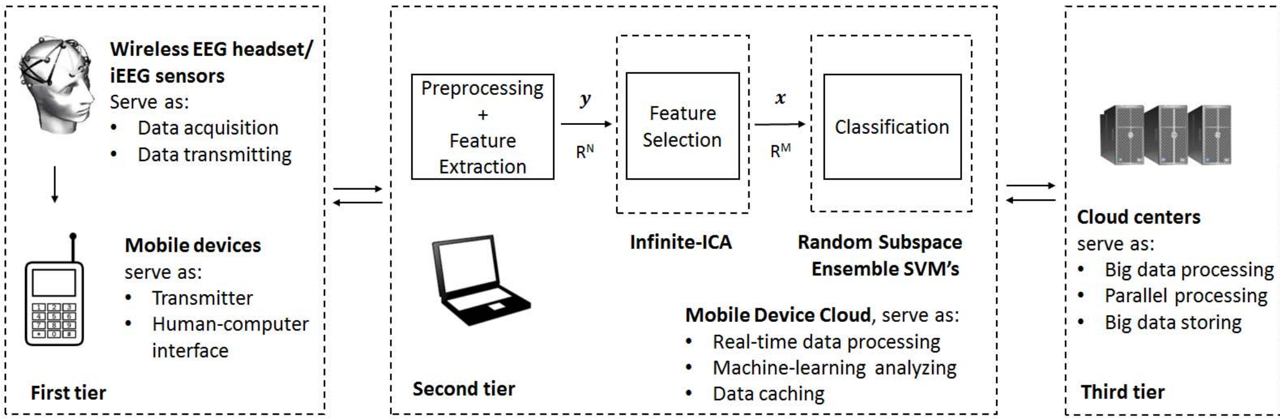

Fig. 2. Conceptual architecture of proposed EEG seizure detection system developed in a multi-tier distributed computing infrastructure and a semantic linked data superstructure. First tier consists of wireless EEG devices and smart phones. Second tier consists of Mobile Device Cloud (MDC) and third tier consists of cloud computing infrastructure. The proposed architecture supports two operation scenarios: (1) Big data analysis using cloud computing paradigm, (2) Interactive and adaptive prediction using real-time brain state and relevant data sets for training and refining brain state prediction.

processed, which is only doable via a cloud-based framework. However, applications of cloud platforms in medicine are still in their infancy and there are only a few studies that have been recently reported. A cloud-based interface to compute cardiac measures using the MapReduce parallel-programming framework is developed in [11] for large volumes of electrophysiological signals. In [12], a pervasive on-line BCI-EEG system using multi-tier fog and cloud computing, semantic linked data search, and adaptive classification models is presented. All in all, existing techniques of feature selection and classification are not suitable for real-time processing and implementation in the cloud because of their execution time, data overfitting, low sensitivity, and high error rate. In contrast, we propose a new method for EEG feature selection and classification that exploits the advantages of cloud computing and to simultaneously increase detection accuracy and decrease error rate.

III. Proposed Work

The proposed system for automatic seizure detection is designed to be performed in the cloud. In this study, a pervasive-computing application for real-time generalized seizure detection is developed, which can be implemented as a cloud-based service. In Sect. III-A, we present an architecture of the BCI-EEG system connected to cloud servers; then, we introduce the seizure-detection system, which consists in performing the following main steps: preprocessing and feature extraction, feature selection by I-ICA, and classification by random ensemble learning. In Sect. III-B, the effects of various sources of noise and artifact are attenuated by filtering. Then, different time and frequency features are extracted for the classification step. In Sect. III-C, we reduce the number of features by I-ICA; this technique increases the classification accuracy by eliminating less-effective features and reduces the computational time by decreasing the number of total features. In Sect. III-D, by random selection of feature subsets, we develop a classification method via ensemble learning that is compatible with high-dimensional data such as iEEG. Then, Majority Voting (MV) is used to aggregate the outputs, achieve consensus, and determine normal vs. epileptic patterns. This new classification method is based on a parallel classification, which makes it appropriate for the big data problem.

A. Conceptual Architecture

The seizure-detection system is proposed as a multi-tier distributed computing structure based on Mobile Device Cloud (MDC) and cloud computing, as shown in Fig. 2. In this architecture, MDC executes tasks in parallel by sharing the workload among multiple nearby mobile devices [13]. The iEEG sensors, dry-electrodes EEG headsets, and smart phones are in the first tier as the interfaces between human and IT technology. An ad-hoc conglomerate of IT devices such as notebooks and home-gateways are in the second tier as MDC for computing purpose. Each MDC server works dually as a data hub and a signal processor. The proposed workflow of signal processing steps—which is shown in Fig. 3 along with its pseudocode in Algorithm 1—extracts time-frequency features from the EEG signals, selects the most informative features, and sends them to the brain state classifiers. These features can also be passed to the next tier for further processing and archiving.

HPC clusters are in the third tier as cloud servers for delivering plenty of computing power, storage capacity, and communication bandwidth to offload the computing burden of the second tier. State-of-the-art techniques for communicating between each tier are Message Queuing Telemetry Transport (QMTT) protocol for interacting between EEG-MDC and RESTful Web Service for interacting between MDC-Cloud. For sending data from the first to the second tier using Bluetooth 4.0 protocol and IEEE 802.11n low-power Wi-Fi technology, a transfer rate of up to $24\ \mathrm{Mbps}$ is supported. With cloud computing, EEG data transport latency through the Internet core runs between 200 and $500\ \mathrm{ms}$ [12].

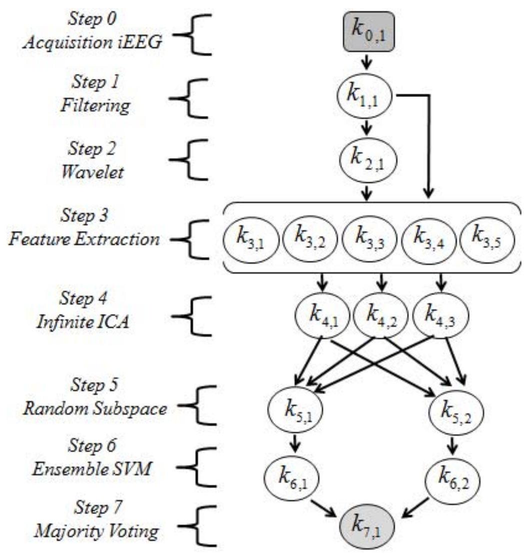

Fig. 3. Workflow of the proposed method consisting of noise and artifact reduction, wavelet, time/frequency-based feature extraction, infinite ICA, ensemble learning by random subspace, ensemble SVM, and majority voting.

Using MDC between cloud and brain sensors delivers two advantages: it provides sub-second real-time responses with minimal communication overhead; and it reduces the amount of traffic between the local area networks and the Internet. As a matter of fact, most of the data sets will be processed and stored in the MDC; therefore, only meta-data needing high computing resources or storages will be uploaded to the cloud.

B. Preprocessing and Feature Extraction

Human internal signals, environmental noise, artifact, and adjacent electrode interference can affect and destroy the information content in iEEG [14]. As first step, some noise and artifacts are removed by filtering. A fourth-order Butterworth bandpass filter is used for cutting signal frequencies; then, to remove unwanted frequencies, a notch filter is applied. Finally, phase distortion is cancelled by using forward and backward filtering. Due to the time-varying nature of iEEG, wavelet transforms can be useful for extracting epileptic spike and capturing the rhythmic nature of seizures; moreover, wavelets have the ability to capture the transient features and to localize them in both time and frequency domains [15]. Then, the outputs of filtering and wavelet transform are used to extract seizure-related features. For this purpose, several time- and frequency-related features have been considered including complexity, mobility, energy, entropy, correlation coefficients, Fast Fourier Transform (FFT), variance, skewness, kurtosis, mean, fractal dimension, frequency band power, peak amplitude, zero crossing, average spectral power, line length, maximal and minimal values, sum absolute value, and some others. Some of these features are defined below.

Hurst Exponent: The iEEG signals including seizure waves consist of various observed signals and independent sources.

Algorithm 1: Proposed Seizure Detection System

Input: Q-dimensional EEG/iEEG Signal
$\mathbf{s} = [s_1, \cdots, s_Q]^T$
Output: classification to seizure (0) or normal (1)
$\pounds \to (0, 1)$
begin
for $i := 1 \to Q$ **do**
  $\mathbf{y} = [y_1, \cdots, y_N]^T$
end
$\mathbf{x} = I - ICA(\mathbf{y})$
for $j := 1 \to \#Subspaces(P)$ **do**
  $r_j = RandomSubspace(\mathbf{x})$
  $v_j = EnsembleSVM(r_j)$
end
for $j := 1 \to P$ **do**
  $\pounds = MV(v_j)$
end
end

We apply ICA to decompose mixed iEEG signals into a set of independent components. Then, we extract the Hurst exponent to find which of the components include epileptic spikes. The range of 0.25–0.45 mV is considered as an epileptic spike [16].

Entropy: The entropy of a signal $\mathbf{s}$, which measures the degree of uncertainty, is given as,

$$H(\mathbf{s}) = -\sum_{i=1}^{Q} s_i \ln(s_i), \tag{1}$$

where i is the index of the samples and Q is the total number of samples per signal segment. We consider the disorder of iEEG segment as a way to find a seizure.

Amplitude: In general, the amplitude of the rhythmic seizure component is greater than in the normal state. Therefore, we extract features based on the signal amplitude, including relative average, coefficient of variation, and duration. To have a normalized measurement, each iEEG epoch is divided by the mean value of the iEEG segment.

Energy: The energy of a signal, given as,

$$E(\mathbf{s}) = \sum_{i=1}^{Q} |s_i|^2. \tag{2}$$

can also be considered as an indicator of a seizure; the higher the energy, the higher the probability of a seizure.

Skewness: The degree of deviation from the symmetry of a Gaussian distribution, which is the third central moment of the amplitude histogram, is measured by the skewness, i.e.,

$$SK(\mathbf{s}) = \frac{\sum_{i=1}^{Q} \frac{(s_i - \bar{s})^3}{N}}{\left(\sum_{i=1}^{Q} \frac{(s_i - \bar{s})^2}{Q-1}\right)^{3/2}}, \tag{3}$$

where $\bar{s}$ is the mean value. For iEEG signals with a symmetrical distribution, the skewness has a nonzero value and it indicates the presence of monophasic events.

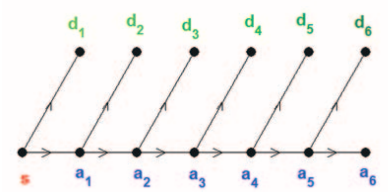

Fig. 4. Wavelet tree showing synthesize signal (s), detail coefficients (d), and approximation coefficients (a).

C. Feature Selection via Infinite ICA

Given a feature space $\mathbf{y} \in \mathbb{R}^N$, feature-selection methods find a mapping $\mathbf{x} = f(\mathbf{y}) : \mathbb{R}^N \rightarrow \mathbb{R}^M$ $(M < N)$ such that the transformed feature vector $\mathbf{x} \in \mathbb{R}^M$ preserves the most information of $\mathbf{y}$. Various feature-selection methods have been proposed to extract a more informative subset of features for classification such as Principal Component Analysis (PCA) and Independent Component Analysis (ICA). These useful statistical feature extractors aim at generating uncorrelated and independent features, respectively. PCA as a decorrelation technique performs a linear mapping of data to a lower-dimensional space in such a way that the variance of the data in the low-dimensional space is maximized.

In general, the correlation matrix of data is constructed and the eigenvectors of this matrix are computed. Then, the eigenvectors corresponding to the largest eigenvalues are used to construct the matrix of PCA transformation. In contrast to PCA, which uses the second-order statistics to convert a set of possibly correlated components into a set of uncorrelated components, ICA makes the mutual information of components equal to zero and depends on all higher-order statistics of $\mathbf{y}$. More precisely, ICA transforms the observed data $\mathbf{y}$ using a linear transformation $\mathbf{G}$ into independent components $\mathbf{x}$ as $\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{e}$, where $\mathbf{e}$ denotes the Gaussian noise [17].

Although PCA and ICA select the feature vector in a lower dimension, they cannot infer the dimensionality of the new-feature vector, i.e., the number of new features must be determined in advance. To solve this problem, we propose to use an extension of ICA, called Infinite ICA (I-ICA) [18], which can also infer the number of independent features automatically from the input data. Previously, this model was successfully applied on various applications such as image denoising, gene expression modeling, image classification and retrieval, and sound source separation [19], [20]; however, this is the first time I-ICA is used in EEG data processing.

For masking on hidden source $\mathbf{x}$, a binary vector $\mathbf{z}$ is determined, where its elements show activity of k^{th} hidden source for the i^{th} data point. This gives us,

$$\mathbf{Y} = \mathbf{G}[\mathbf{X} \odot \mathbf{Z}] + \mathbf{E}, \tag{4}$$

where $\mathbf{Y}, \mathbf{X}, \mathbf{Z}, \mathbf{E}$ denote the concatenation of $\{\mathbf{y}_i\}_{i=1}^N$, $\{\mathbf{x}_i\}_{i=1}^N$, $\{\mathbf{z}_i\}_{i=1}^N$, and $\{\mathbf{e}_i\}_{i=1}^N$, respectively; and $\odot$ denotes the element wise multiplication. Since $\mathbf{Z}$ has infinitely many rows, an infinite number of hidden sources can be achieved.

Algorithm 2: Feature Selection by I-ICA

Input: N-dimensional features $\mathbf{y} = [y_1, \cdots, y_N]^T$
Output: M-dimensional Selected features $\mathbf{x} = [x_1, \cdots, x_M]^T$
begin
for $i := 1 \rightarrow N$ **do**
  $Y = concatenation\{\mathbf{y}_i\}_{i=1}^N$
end
Calculate $p(\mathbf{Z}|\pi_1, ..., \pi_K)$ by (5)
for $k := 1 \rightarrow K$ **do**
  for $i := 1 \rightarrow N$ **do**
    Find $\mu_\pm$ and σ^2 and e_{ki}° by (7)
    Find $p(x_k|G, x_{-ki}, y_i, z_i)$ by (6)
  end
end
Find X by (4)
Extract $\mathbf{x}$ from X
end

For N data points and K hidden sources, the distribution of matrix $\mathbf{Z}$ is defined by,

$$p(\mathbf{Z}|\pi_1, ..., \pi_K) = \prod_{k=1}^{K}\prod_{i=1}^{N} P(z_{ki}|\pi_k) = \prod_{k=1}^{K} \pi_k^{m_k}(1-\pi_k)^{N-m_k}, \tag{5}$$

where z_{ki} indicates activity of k^{th} source for i^{th} sample using probability of π_k and $m_k = \sum_{i=1}^N z_{ki}$ indicates the total number of active sources. To define $\mathbf{E}$, a Gaussian noise has been considered with variance σ_e^2. Finally, for inferring $\mathbf{X}$ hidden sources from $\mathbf{Y}$ observed data using the $\mathbf{G}$ mixing matrix (which has Z active sources), an inference is defined. Gibbs sampling is used to sample elements with $z_{ki} = 1$. This sampling proceeds by sampling successively from the conditional distribution of one parameter given all others by Baye's rule [18]. Therefore, the result is a piecewise Gaussian distribution that is defined by,

$$\mathbf{p}(x_{ki}|\mathbf{G}, x_{-ki}, y_i, z_i) = \begin{cases} \mathcal{N}(x_{ki}; \mu_-, \sigma^2) & if \quad x_{ki} > 0 \\ \mathcal{N}(x_{ki}; \mu_+, \sigma^2) & if \quad x_{ki} < 0, \end{cases} \tag{6}$$

where

$$\mu_\pm = \frac{g_k^T e_{ki}^\circ \pm \sigma_e^2}{g_k^T g_k}, \quad \sigma^2 = \frac{\sigma_e^2}{g_k^T g_k}, \quad e_{ki}^\circ = (e_{ki}|z_{ki} = 0), \tag{7}$$

and $\mathbf{g}_k$ is the k^{th} column of $\mathbf{G}$.

The proposed method for feature selection is described in Algorithm 2. In Sect. IV, we discuss how the algorithm successfully extracts independent components of EEG features and is also able to indicate when the sources are active. The I-ICA characteristics make it a good candidate for feature selection in big data problems. Moreover, by reducing the number of features this technique makes the classification task less complex and therefore faster, which is a requirement for real-time classification.

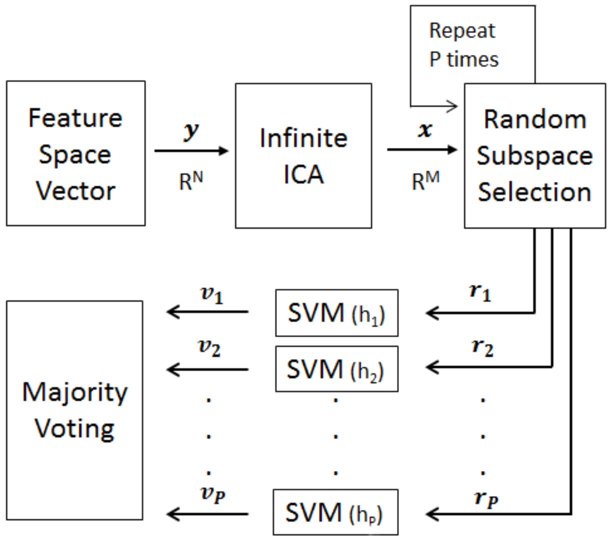

Fig. 5. Random Subspace ensemble with SVM as the base classifiers using I-ICA for feature selection. The feature space is $\mathbf{y} \in \mathbb{R}^N$; the feature selection method (I-ICA) finds a mapping $\mathbf{x} = f(\mathbf{y})$: $\mathbb{R}^N \to \mathbb{R}^M$ $(M < N)$. Random Subspace utilizes random subsets in the feature space and makes P subspaces denoted by $\mathbf{r}_1, ..., \mathbf{r}_P$. Then, SVM's classifiers are used on each subspace to classify the input data as healthy or epileptic. Finally, using Majority Voting (MV), the output with the highest number of votes is chosen.

D. Random Subspace Ensemble Classification

Most of the existing iEEG works use only a small training dataset, notwithstanding the inherent high dimensionality of the problem, which makes features-to-instance ratio extremely high. It has been shown that training a classifier by a small training dataset with high dimensionality leads to low classification performance due to overfitting [21]. To address this problem and also reduce the computational load, which is necessary for real-time and pervasive analysis, we propose a new classification technique, namely a Random Subspace ensemble method with SVM as the base classifiers (see Fig. 5).

To avoid overfitting caused by high dimensionality and small size data, ensemble learning techniques have been considered as alternative to single classifiers. These techniques combine weak classifiers and aggregate multiple learning algorithms to provide a collective decision and improve overall learning performance. Random subspace ensemble method uses a random sample of features instead of using all the features. Unlike other ensemble techniques such as Bagging and AdaBoost [22], which utilize random subsets of the samples in the input training space, Random Subspace utilizes random subsets in the feature space. In our solution, decision rules in each SVM classifier are learned with a randomly-selected feature subspace using all the training samples. The M dimensional output of I-ICA is randomly divided into P subspaces denoted by $\mathbf{r}_1, ..., \mathbf{r}_P$.

After preparing the feature subspaces, SVMs as an efficient classifier are used on each subspace to classify the input data as normal or epileptic. A nonlinear function $\Phi(\mathbf{r})$ is used to map input space into a higher-dimensional feature space that

Algorithm 3: Random Subspace Ensemble Classification

Input: M-dimensional Selected features,
  $\mathbf{x} = [x_1, \cdots, x_M]^T$
Output: Classification result as seizure (0) or normal (1),
  $z \to (0, +1)$
begin
for $i := 1 \to$*#subjects* **do**
  Find train_input by eliminate i-th column of X
  Find test_input using i-th column of X
  Find train_target by eliminate i-th column of $target$
  Find target_test using i-th column of $target$
  for $i := 1 \to$*#SVMs(P)* **do**
    Find train_features as $RandomSubspace$ of train_input
    Find $k(r_i, r_j)$ by (10)
    Calculate quadratic optimization for $w(\alpha)$
    Perform SVM training by (train_input, train_target, $k(r_i, r_j)$)
    Find $\mathbf{v_j}$ by train_features by (8)
  end
  Find final_output by MV of $\mathbf{v_j}$
end

corresponds to a linear surface in the feature space. For each classifier, the decision function v is found by,

$$v = sgn(\mathbf{w}.\Phi(\mathbf{r}) - b), \tag{8}$$

where b is a bias and sgn is the sign of a real number. Then, the classification problem is defined as distinguishing normal against seizure features. This is done by identifying an hyperplane $\frac{2}{\|\mathbf{w}\|}$ that divides the features of two classes, which corresponds to the problem of maximizing a quadratic function of defined variables subject to their linear constraints. Such quadratic programming optimization is solved by,

$$\mathbf{w}(\alpha) = 0.5\,\|\mathbf{w}\|^2 - \alpha[\mathbf{o}(\mathbf{w}.\Phi(\mathbf{r}) - b) - 1], \tag{9}$$

where α is defined as a set of Lagrange multipliers.

Since most of the time iEEG training sets are not separable using a linear function, a Gaussian Radial Basis Function (GRBF) kernel is used to minimize the classification error [23]. This kernel function is defined as,

$$\mathbf{k}(\mathbf{r}_i, \mathbf{r}_j) = \Phi^T(\mathbf{r}_i).\Phi(\mathbf{r}_j) = \exp\{\|\mathbf{r}_i - \mathbf{r}_j\|_2^2/2\sigma^2\}, \tag{10}$$

where $\mathbf{r}_i$ and $\mathbf{r}_j$ are two feature vectors, and $\|\mathbf{r}_i - \mathbf{r}_j\|_2^2$ is the squared Euclidean distance between $\mathbf{r}_i$ and $\mathbf{r}_j$. Finally, as a standard rule for combining the results of all classifiers, Majority Voting (MV) is used. By the use of MV, the output with the highest number of votes is considered as the final output of the overall system. Since the strength of a correlation does not necessarily predict the outcome for a new observation, cross-validation technique is used for increasing the generality, which can assess how the results of a statistical analysis can be generalized to an independent data set. We used leave-one-out as an exhaustive cross validation for directly estimating the

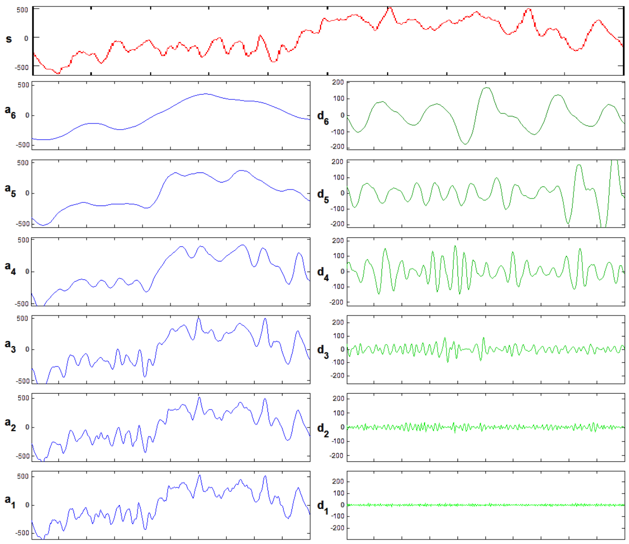

Fig. 6. Synthesized signal for 1-second seizure data segment: S, 6 levels of approximation coefficient: a1-a6 (Left), 6 levels of detail coefficients: d1-d6 (Right). The x-axis represents samples and the y-axis shows the signal in μV. Scale 4-6 of detail coefficients represent seizure activity.

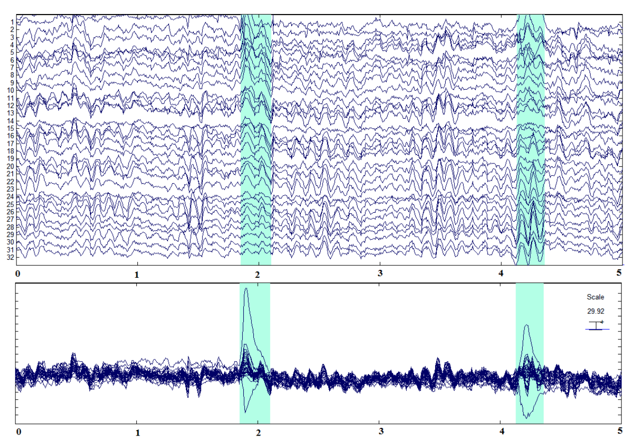

Fig. 7. EEG segments of 32 channels showing epileptiform normal variants in a few channels (top) and its stack EEG segments (bottom) in highlighted areas. The horizontal line is Time [s] and the vertical line is Amplitude [μV].

predictive accuracy of a particular statistical model. For this purpose, the model is fitted to subsets of data and the accuracy of the model is found with held-out sample. The pseudocode of the proposed classification method is shown in Algorithm 3.

IV. Performance Evaluation

A proof-of-concept prototype of the proposed pervasive EEG-based seizure detector was developed, where we used a benchmark epileptic dataset for the first tier; an HP laptop with Intel i5 processor, 8 GB RAM, and a 4.4 Ah battery for the second tier; and a supercluster of computers hosted by Amazon Elastic Compute Cloud (EC2)[1] as the third tier.

TABLE I
Source of variability, Sum of Squares (SS) of each source, degrees of freedom (df) of a source, Mean Square (MS) as the ratio SS/df, ratio of the mean squares as F-statistic, and p-value derived from the Cumulative Distribution Function (CDF) of F are shown as the standard ANOVA.

Group	Source	SS	df	MS	F-statistic	p-value
	Columns	2.89e+09	8	3.62e+08	11.45	1.85e-14
Normal	Error	1.11e+10	351	3.16e+07		
	Total	1.39e+10	359			
	Columns	1.96e+10	8	2.45e+09	21.92	4.64e-27
Seizure	Error	3.92e+10	351	1.11e+08		
	Total	5.88e+10	359			

The clinical iEEG dataset of eight patients (52 normal and 52 seizure segments) with temporal and extratemporal lobe epilepsy has been used, which was jointly developed by the U. of Pennsylvania and the Mayo Clinic, and sponsored by the American Epilepsy Society [24]. The iEEGs are recorded in depth electrodes implanted along anterior-posterior axis of the hippocampus, and in subdural electrode grids in various locations.[2] Seizure data segments (labeled ictal) and non-seizure data segments (labeled interictal) by sampling rate from 500 to $5,000$ Hz have been used for training and testing. The entire seizures were recorded in the ictal data segments and all data are organized into 1 s EEG clips. The mean seizure duration for each subject is covered in the interictal data segments with the restriction of no less than one hour before or after a seizure.

As first step, various sources of noise and artifacts are attenuated via filtering, where a fourth-order Butterworth bandpass filter (0.5-150 Hz) is used for cutting frequencies. Then, to remove some unwanted frequency of oscillator, a notch filter set at 50 Hz is applied. In the next step, the phase distortion is cancelled by using forward and backward filtering. The appropriate wavelet and level of decomposition are chosen based on the input signal and application. Based on our evaluation of common wavelets, we use Daubechies 4 (db4) to find approximation and detail for iEEG data. Since seizure activities at iEEGs commonly occur in 3-25 Hz, the detail coefficients have been investigated to find this frequency range. First, we consider sampling frequency of the data (500 Hz). By the Nyqvist criteria, the maximum frequency of data is obtained at 250 Hz. Finally, by coefficient representation in each scale, the frequency range of 3-30 Hz is covered in scales of 4, 5, and 6. The 6 levels of approximation coefficient (a1-a6) and the corresponding detail coefficients (d1-d6) are shown in Fig. 6. The seizure is covered in d4-d6 scales.

Many EEG patterns that resemble epileptogenic abnormalities are not associated with epilepsy or any neurologic conditions such as small sharp spikes, wicket spikes, phantom

[1] Amazon EC2 is a web service that provides resizable compute capacity in the cloud; its web-service interface allows users to obtain and configure capacity, which provides a complete control of the computing resource.

[2] Hybrid depth and subdural electrodes contain clinical macroelectrodes and additional microwire arrays; they are manufactured by Adtech Medical Instrument Corporation, Racine, WI and by PMT Chanhassen, MN.

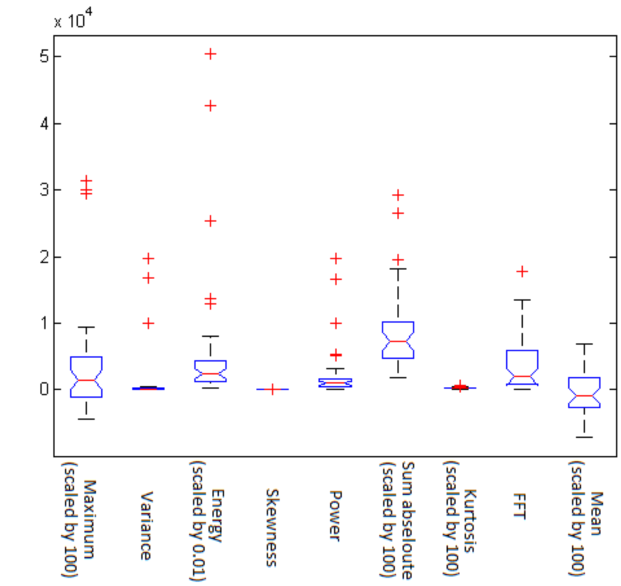

Fig. 8. Graphical depiction of some metrics on the extracted features for normal group. Amplitude maximum (multiplied by scaled factor 100), variance, energy (multiplied by scaled factor 0.01), skewness, power, sum of absolute value (multiplied by scaled factor 100), FFT, and mean value (multiplied by scaled factor 0.01) are shown, respectively. Outliers are plotted by plus signs.

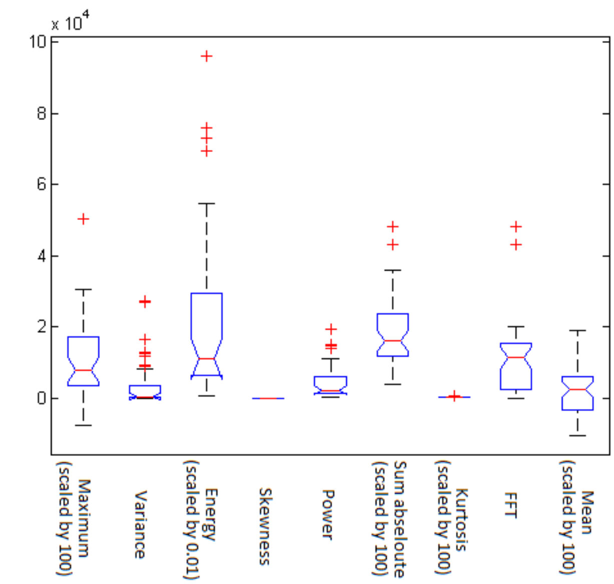

Fig. 9. Graphical depiction of some metrics on the extracted features for the seizure group. Maximum amplitude (multiplied by scaled factor 100), variance, energy (multiplied by scaled factor 0.01), skewness, power, sum of absolute value (multiplied by scaled factor 100), FFT, and mean value (multiplied by scaled factor 0.01) are shown, respectively. Outliers are plotted by plus signs.

waves, and paroxysmal rhythmic discharges [25]. Since such patterns have no clinical significance for seizure detection, they are called epileptiform normal variants. These patterns are one of the major causes of false seizure detection in automatic methods. However, their recognition is important for avoiding over-interpretation. After analyzing this type of patterns, we remove the features that resemble those patterns (see Fig. 7) using frequency filtering followed by block scaling of amplitude and slope [26].

Table I displays the results of a standard ANOVA analysis [1], while Figs. 8 and 9 depict some metrics on the extracted features for normal and seizure subjects by quartile. The large F-statistic and small value of p in Table I correspond to a large difference in the center lines of the box plots between Figs. 8 and 9. In the proposed ensemble method, we use 64 features as the input parameters of SVMs. In this study, by random selection of features in each subset, we find five subsets of features. Then, features in each subset are used to train an SVM. Such randomness combination of multiple features provides a better classification accuracy compared to using all of the features.

TABLE II
ACCURACY, SENSITIVITY, SPECIFICITY, FPR, AND FNR FOR PROPOSED CLASSIFICATION COMPARED WITH THE OTHER METHODS

Methods	Accuracy	Sensitivity	Specificity	FPR	FNR
Proposed Method	0.95	0.96	0.94	0.06	0.04
Proposed Method (linear)	0.91	0.92	0.90	0.10	0.08
MLP Neural Network	0.82	0.81	0.83	0.17	0.19
Linear SVM	0.84	0.83	0.85	0.15	0.17
Non-linear SVM	0.85	0.85	0.87	0.13	0.15

Figures 10 and 11 show the result of linear and nonlinear classification, respectively, for one of the five random subset of features. We define the following performance metrics to evaluate and compare our results against previous methods. Accuracy measures the proportion of both seizure and normal signals that are correctly identified as epileptic and healthy, respectively. Sensitivity (true positive rate) measures the proportion of seizure signals that are correctly identified as seizures. Specificity (true negative rate) measures the proportion of normal signals that are correctly identified as normal. False Positive Ratio (FPR) measures the proportion of normal signals that are incorrectly identified as seizures. Finally, False Negative Ratio (FNR) measures the proportion of seizure signals that are correctly identified as normal [27]. Using the leave-one-out cross-validation approach, the performance results are as follows: accuracy=0.95, sensitivity=0.96, specificity=0.94, FPR=0.06 , and FNR=0.04. The classification results of our method with non-linear and linear base classifiers are compared to single SVM and Multi-Layer Perception (MLP) Neural Network in Table II. The neural-network structure is designed in three layers using the Levenberg-Marquardt optimization for the training phase. Experimental results in Table II show that the random subspace method outperforms previous methods for a subset of features selected by our proposed method.

V. DISCUSSION AND FUTURE WORK

There is a crucial need to develop new methods using advanced technologies such as cloud and mobile computing in order to assist in the processing of EEG data and develop

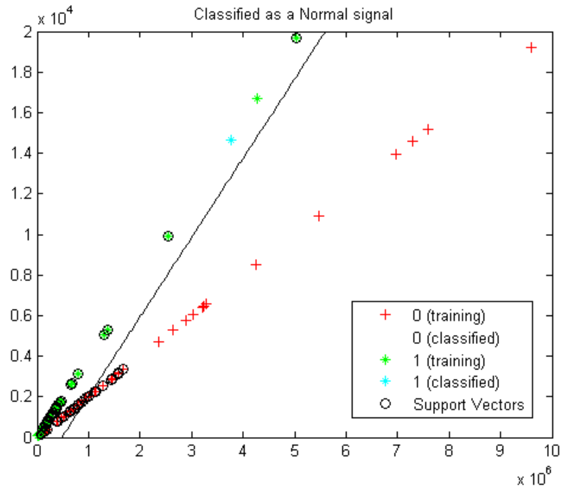

Fig. 10. Results of linear classification in one subset of features (P1, P2), which are made with I-ICA. Red signs show features extracted from the seizure group, green signs show features extracted from the normal group, and the blue sign is a test signal, which is originally a normal brain activity that is correctly identified and classified as normal by the classifier.

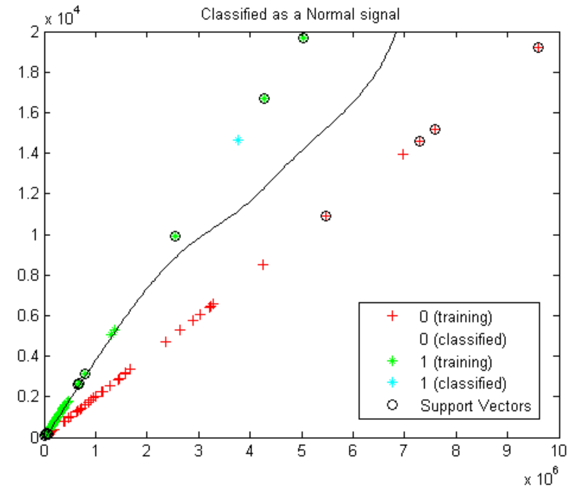

Fig. 11. Results of non-linear classification in one subset of features (P1, P2), similarly to Fig. 10.

pervasive computing applications such as real-time seizure detection. Our system's service platform can be hosted at a remote location and may be used anywhere via Internet. Instead of installing the seizure-detection program on multiple computers, a web-based platform can be of support in such a way as to guarantee greater accessibility, higher security, and little financial cost. The service model is designed as Platform as a Service (PaaS) and the computing platform is delivered in a cloud provider. The general structure of the cloud framework is divided into three parts—consisting of the transmission, computing, and searching among the history patients—and the seizure detection component is offered as a cloud service in the PaaS. Users of the system can access the data and seizure-detected segments without worrying about the underlying computation and storage services, which are offered in a transparent way to the users. The physical storage provides in-situ and real-time access to the original and processed data. The cost and complexity of buying, installing, updating, and managing the underlying software layer is reduced and the users (physicians, nurses, patients) do not have to allocate resources manually. To study the feasibility of collaboration within the multi-tier architecture (local and cloud servers), the network latency offered by different servers in Amazon EC2 was analyzed and the Round Trip Time (RTT) for servers located at different geographical locations (Virginia, Oregon, Singapore, and Ireland) was obtained. The RTT of 64B EEG segments, which repeatedly have been sent from the first to the third tier, is calculated at 10 days using the "ping" command. The mean time is reported in Fig. 12 over different day times.

In sum, the proposed pervasive framework can be implemented as a cloud-based service and it has the advantages of cloud computing such as running detection algorithms for multiple users simultaneously and the aggregation of data. Moreover, the proposed feature selection and classification provide a possibility of seizure detection of massive scaling, which helps better train the classifier. Finally, computing is not limited to the computational power of local MDC and more complex computations can be performed in the cloud for enhanced decision making.

As future work, we will consider additional classification methods to improve the detection rate, e.g., random forest and deep learning in the training phase of our ensemble classifier. We aim at extending our framework to use historical medical records of patients in order to make more informed decisions. We also plan to study other neurological disorders besides epilepsy such as sleep disorders, coma, encephalopathies, and brain necrosis to develop a complete neurological disorder system. This system will use other vital biosignals besides EEG and will require processing a variety of models to estimate different neurological diseases in real time. In parallel, we will complete the autonomic loop by considering the responsive neurostimulation signal. We plan to work on applying an appropriate stimulation signal to remove the seizures as the second stage of the autonomic computing system for managing the human brain of epileptic patients. The success of the therapeutic process is dependent on adequate sensing of seizures and stimulation algorithms as well as on a fast coupling between the two [5]. We plan to conduct more experiments in order to understand how the estimation of EEG variance can be used as feedback for responsive neurostimulation so to implement the second stage of autonomic computing system and regulate the human brain. A reasonable approach initially might be to send a notification to physicians along with summarized explanatory information that would allow them to make their own judgment. Also, a link that they could click on to apply the stimulation signal or to not apply it, which could also be used to help further train the system.

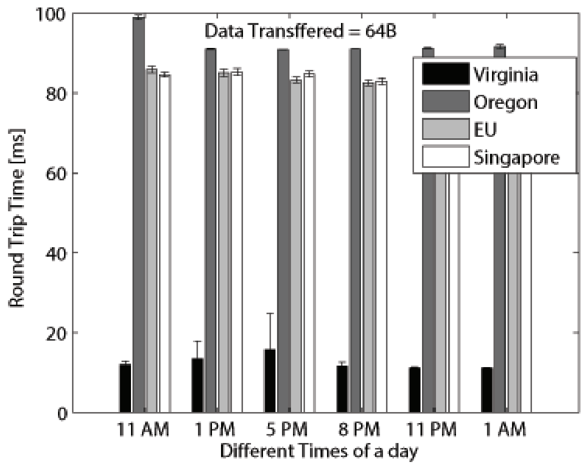

Fig. 12. Round Trip Time (RTT) between Amazon EC2 servers (third tier) and a local machine (first tier) for EEG clips. At any time, the lowest RTT is observed for the server located in Virginia (the closest to our location).

ACKNOWLEDGMENTS

We thank Hamid Soltanian-Zadeh, Ph.D., and Mohammad Nazem-Zadeh, Ph.D., Dept. of Radiology and Research Administration, Henry Ford Health System, MI, USA for helping analyze brain signals and for referring epileptic cases. We are grateful to Kost V. Ellisevich, M.D., Dept. of Clinical Neuroscience, Spectrum Health System, MI, USA for his collaboration. Finally, we are in debt with the anonymous reviewers of the conference for their valuable comments, which helped us improve the paper.

REFERENCES

[1] M.-P. Hosseini, M.-R. Nazem-Zadeh, D. Pompili, K. Jafari-Khouzani, K. Elisevich, and H. Soltanian-Zadeh, "Comparative performance evaluation of automated segmentation methods of hippocampus from magnetic resonance images of temporal lobe epilepsy patients," *Medical physics*, vol. 43, no. 1, pp. 538–553, 2016.

[2] D. E. Blum, J. Eskola, J. J. Bortz, and R. S. Fisher, "Patient awareness of seizures," *Neurology*, vol. 47, no. 1, pp. 260–264, 1996.

[3] A. G. Correa, L. Orosco, P. Diez, and E. Laciar, "Automatic detection of epileptic seizures in long-term eeg records," *Computers in Biology and Medicine (Elsevier)*, 2014.

[4] M.-P. Hosseini, H. Soltanian-Zadeh, S. Akhlaghpoor, A. Jalali, and M. Bakhshayesh Karam, "Designing a new cad system for pulmonary nodule detection in high resolution computed tomography (hrct) images," *Tehran University Medical Journal (TUMJ)*, vol. 70, no. 4, pp. 250–256, 2012.

[5] S. Carrette, P. Boon, M. Sprengers, R. Raedt, and K. Vonck, "Responsive neurostimulation in epilepsy," *Expert review of neurotherapeutics*, vol. 15, no. 12, pp. 1445–1454, 2015.

[6] D. Kuzum, H. Takano, E. Shim, J. C. Reed, H. Juul, A. G. Richardson, J. de Vries, H. Bink, M. A. Dichter, T. H. Lucas, D. A. Coulter, E. Cubukcu, and B. Litt, "Transparent and flexible low noise graphene electrodes for simultaneous electrophysiology and neuroimaging," *Nature communications*, vol. 5, 2014.

[7] Y. U. Khan and J. Gotman, "Wavelet based automatic seizure detection in intracerebral electroencephalogram," *Clinical Neurophysiology (Elsevier)*, vol. 114, no. 5, pp. 898–908, 2003.

[8] S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, "Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 9, pp. 1545–1551, 2007.

[9] R. Hopfeng, B. S. Kasper, W. Graf, S. Gollwitzer, G. Kreiselmeyer, H. Stefan, and H. Hamer, "Automatic seizure detection in long-term scalp eeg using an adaptive thresholding technique: A validation study for clinical routine," *Clinical Neurophysiology (Elsevier)*, vol. 125, no. 7, pp. 1346–1352, 2014.

[10] R. Aghazadeh, P. Shahabi, J. Frounchi, and M. Sadighi, "An autonomous real-time single-channel detection of absence seizures in wag/rij rats," *General physiology and biophysics*, 2015.

[11] S. S. Sahoo, C. Jayapandian, G. Garg, F. Kaffashi, S. Chung, A. Bozorgi, C. Chen, K. Loparo, S. D. Lhatoo, and G. Zhang, "Heart beats in the cloud: distributed analysis of electrophysiological 'big data' using cloud computing for epilepsy clinical research," *Journal of the American Medical Informatics Association*, vol. 21.

[12] J. K. Zao, T. Gan, C. You, C. Chung, Y. Wang, S. J. Mndez, T. Mullen, C. Yu, C. Kothe, C. Hsiao, and al., "Pervasive brain monitoring and data sharing based on multi-tier distributed computing and linked data technology," *Frontiers in human neuroscience*, vol. 8, 2014.

[13] H. Viswanathan, E. K. Lee, I. Rodero, and D. Pompili, "Uncertainty-aware autonomic resource provisioning for mobile cloud computing," *IEEE Trans. on Parallel and Distributed Systems (TPDS)*, vol. 26, no. 8, pp. 2363–2372, Aug. 2015.

[14] M.-P. Hosseini, M. R. Nazem-Zadeh, D. Pompili, and H. Soltanian-Zadeh, "Statistical validation of automatic methods for hippocampus segmentation in mr images of epileptic patients," in *Proc. of IEEE International Conference of the Engineering in Medicine and Biology Society (EMBC)*, 2014, pp. 4707–4710.

[15] S. Mercy, "Performance analysis of epileptic seizure detection using dwt & ica with neural networks," *International Journal of Computational Engineering Research*, vol. 2, no. 4, pp. 1109–1113, 2012.

[16] N. Arunkumar, V. S. Balaji, S. Ramesh, S. Natarajan, V. R. Likhita, and S. Sundari, "Automatic detection of epileptic seizures using independent component analysis algorithm," in *Proc. of IEEE International Conference of Advances in Engineering, Science and Management (ICAESM)*, 2012, pp. 542–544.

[17] M. Rahmati, P. Pandey, and D. Pompili, "Separation and classification of underwater acoustic sources," in *Proc. of IEEE Underwater Communications and Networking (UComms)*, 2014.

[18] D. Knowles and Z. Ghahramani, "Infinite sparse factor analysis and infinite independent components analysis," in *Independent Component Analysis and Signal Separation*. Springer Berlin Heidelberg, 2007, pp. 381–388.

[19] B. Gholami and A. Hajisami, "Kernel auto-encoder for semi-supervised hashing," *Proc. of IEEE Winter Conference on Applications of Computer Vision*, 2016.

[20] T. Otsuka, K. Ishiguro, H. Sawada, and H. G. Okuno, "Bayesian unification of sound source localization and separation with permutation resolution." in *AAAI*, 2012.

[21] A. Leiva-Murillo, J. M.and Arts-Rodrguez, "Maximization of mutual information for supervised linear feature extraction," *IEEE Transactions on Neural Networks*, vol. 18, no. 5, pp. 1433–1441, 2007.

[22] T. G. Dietterich, "Ensemble methods in machine learning," in *Multiple classifier systems*. Springer, 2000, pp. 1–15.

[23] M. P. Hosseini, M. R. Nazem-Zadeh, F. Mahmoudi, H. Ying, and H. Soltanian-Zadeh, "Support vector machine with nonlinear-kernel optimization for lateralization of epileptogenic hippocampus in mr images," in *Proc. of IEEE International Conference of the Engineering in Medicine and Biology Society (EMBC)*, 2014, pp. 1047–1050.

[24] M. Stead, M. Bower, B. H. Brinkmann, K. Lee, W. Marsh, F. B. Meyer, B. Litt, G. Van, and G. Worrell, "Microseizures and the spatiotemporal scales of human partial epilepsy," *Brain*, p. awq190, 2010.

[25] T. A. Pedley, "Eeg patterns that mimic epileptiform discharges but have no association with seizures," *Current Clinical Neurophysiology: Update on EEG and Evoked potentials. New York: Elsevier/North Holland*, 1980.

[26] D. T. Barkmeier, A. K. Shah, D. Flanagan, M. D. Atkinson, R. Agarwal, D. R. Fuerst, K. Jafari-Khouzani, and J. A. Loeb, "High inter-reviewer variability of spike detection on intracranial eeg addressed by an automated multi-channel algorithm," *Clinical Neurophysiology (Elsevier)*, vol. 123, no. 6, pp. 1088–1095, 2012.

[27] M.-R. Nazem-Zadeh, J. M. Schwalb, H. Bagher-Ebadian, F. Mahmoudi, M.-P. Hosseini, K. Jafari-Khouzani, K. V. Elisevich, and H. Soltanian-Zadeh, "Lateralization of temporal lobe epilepsy by imaging-based response-driven multinomial multivariate models," in *Proc. of IEEE International Conference of Engineering in Medicine and Biology Society (EMBC)*, 2014, pp. 5595–5598.