



Extracting Discriminative Features in Reactive Power Variations by 1-Persistence Parallel Fragmented Hybrid Feature Selection Scheme for Transient Stability Prediction

Seyed Alireza Bashiri Mosavi^{1*}

¹*Imam Khomeini International University – Buin Zahra Higher Education Center of Engineering and Technology,
Iran*

* Corresponding author's Email: abashirimosavi@bzte.ac.ir

Abstract: The corrective action for power systems over transient space requires the transient stability prediction (TSP) in an accurate and timely manner. To this end, applying data mining techniques to achieve high performance on TSP is an inevitable approach. One of the influential factors on TSP performance is extracting the most relevant features (MRFs) of transient data. In this paper, first, we constructed the transient dataset in the form of reactive power-based two-variate time series (RP2vTS). Next, for exploiting the optimal feature subset, the MRFs of RP2vTS are obtained by the 1-persistence parallel fragmented hybrid feature selection scheme (PFHFS) based on filter and wrapper methods. Finally, to evaluate the efficacy of the proposed framework, experimental comparisons on inter-connected New England and New York systems (NETS-NYPS) were applied using the support vector machine (SVM) classifier. The results showed that the proposed framework by selecting MRFs of RP2vTS offers high-performance capacity on TSP.

Keywords: Most relevant features (MRFs), Transient stability assessment (TSA), Hybrid feature selection.

1. Introduction

The power system vulnerability is a constant concern with network managers in exploitation conditions on generation, transmission, and distribution power lines. In fact, reliability is the unique goal in the design and development of power systems toward the secure electric power supply. Hence, continuous monitoring in the form of the situational awareness dashboard for observing feature variability of power system components is a necessary scenario to satisfy the power system security index. The pivot problem in grid security is power system stability. One of the most important branches of power system stability is transient stability which is focused on interoperability of the power system components when severe or sudden fault occurrence [1]. In this regard, the confluence of phasor measurement units (PMUs) and data mining (DM) technologies as new paradigms are significant

hand tools for power system operators (PSOs) on transient stability assessment (TSA) [2]. Nowadays, installed PMUs equip as part of an interconnected transmission grid is the basic factor for precise measuring of power system component features (e.g.; rotor angle dynamics) on TSA [3]. Furthermore, DM technology that interlaced statistics, machine learning, and artificial intelligence play a supplementary role to discover knowledge on transient variation gained by PMUs [4-8]. Along with path-breaking technologies (DM and PMUs), it is inevitable to compromise between the triple issues namely high dimensional multivariate transient data, the processing time of transient analysis, and accuracy of TSP. Hence, the well-suited strategy for addressing the triangulated challenge is to select the optimal feature set. Therefore, this solution can be vetted from the following two aspects:

(a) Defining relevant feature against transient analysis: The multivariate time series obtained by PMUs per feature of the power system component

causes gathering the huge volume of transient data points. Applying such high-dimensional transient data as input to DM techniques increases the computational complexity to find the separator hyperplane in the transient feature space (TFS). One of the significant issues which affect the performance of the learning model is the selection of proper features on transient analysis. The discussion on reasons for choosing effective features for transient stability prediction (TSP) is an important task. In fact, this issue will make a significant contribution to reducing TFS and computational complexity on TSP. In transient studies, several features have been considered by scholars. According to [9-11] the rotor angle monitoring has been considered for transient stability status prediction in the form of a PMU-based model-free approach. Also in [12-14], transient voltage and transient frequency are considered as predictors for TSA. The active power variations are considered for TSA on multi-terminal systems in Reference [15,16].

(b) Applying feature selection scheme (FSS): Although defining relevant features on transient analysis cause to reducing TFS (above-aspect), it is not entirely efficient. In fact, due to the occurrence of swift-snap events is a dominant property in transient stability, the corrective control action on the power system should be launched in a timely manner [2]. Hence timely actions by PSOs have a direct relationship with the processing time of TSP which itself depends on TFS (transitive relation). For this purpose, a great option to respond to the time-oriented bottleneck on TSA is the feature selection process which one of the interdisciplinary approaches in DM technology. In transient studies, supervised feature selection techniques (wrapper and filter methods) to remove irrelevant features have been considered by scholars. According to [17], the most relevant features of voltage magnitude or rotor angle based on filter FSS (Relief) have been considered for TSP. The Relief FSS to select optimal features for diagnosing rotor faults has been considered in [18]. The min-redundancy and max-relevance (mRMR) FSS applied on the set of inter-complementary dynamic stability features for TSA [19]. In Reference [20], the FCBF as feature pre-screening has been considered for the total transfer capability calculation model regarding transient stability constraint.

Previous researches for TSA which is concerning defining relevant feature against transient analysis and applying feature selection scheme (See Section 1; Paragraph (a) & (b)), can be developed by responding to the following challenges:

1) Specific focus on features that are compatible with real conditions of the power grid is an important

issue on real-time TSA. The literature review on transient analysis shows that the TSP of the power system is restricted to the measurement of a quantity such as a rotor angle, active power, or voltage. For example, the result obtained from the TSP of the power system in the presence of the rotor angle does not correspond to the result obtained in the real environment in terms of the same feature. In fact, according to the nature of the transient instability, we must define the novel generalized feature that represents the cross-effects of parameters of power system components in the transient state. This did not occur in previous research with respect to (a) aspect. Hence, this issue motivated us to consider this challenge.

2) Focus on the type of the transient feature space in the form of the optimal feature selection scheme is a significant issue. In previous transient studies, presenting FSS on the point-based feature space has been considered by scholars. In fact, the designing framework for FSS in the presence of the multivariate transient feature space has remained the motivational challenge in terms of (b) aspect. On the other hand, in the previous FSS-based transient studies, statistical and machine learning-based techniques (namely, filter and wrapper method) are applied cohesively (on the whole of feature space) caused to loss of many useful features. In fact, extracting intrinsic information by performing the segment-based filter/wrapper approach leads to the survival of the optimal-blurred features.

According to what was mentioned above, defining the generalized feature and selecting the most discriminative features over transient sequences to achieve high performance (processing time and accuracy) on TSA is defined as the motivational issues in this paper. To this end, first, we generated the transient dataset in the form of reactive power two-variate time series (RP2vTS). Next, the 1-persistent parallel fragmented hybrid feature selection scheme (PFHFS) is offered for selecting the most relevant features (MRFs) per univariate time series. In the proposed 1-persistence PFHFS, we applied the filter and wrapper phase which is derived from the information theory concept and performance retrieved from employing the classification technique, respectively. Also, based on obtained MRFs for each univariate time series, the proposed PFHFS applied on MRFs once again (1-persistence scenario). Finally, the effectiveness of selected final optimal features on TSA is evaluated via the SVM-cross validation procedure.

The rest of the paper is organized as follows: In

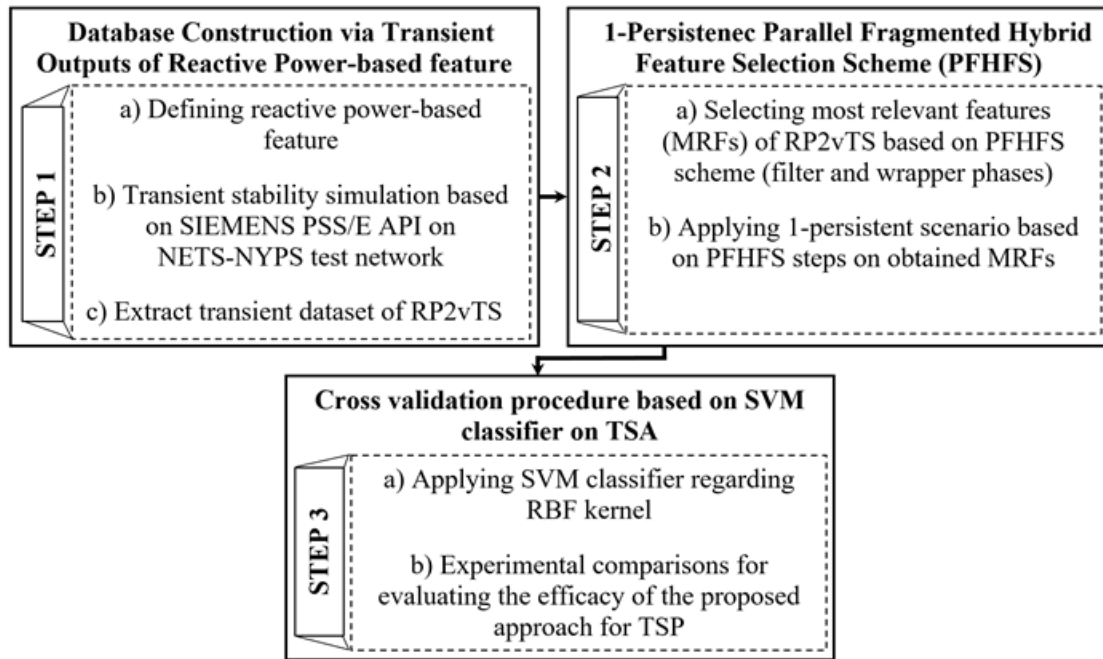


Figure. 1 The proposed framework for transient stability prediction

section 2, first, we define two-variate time series via transient reactive power-based measurement and then offer a detailed description of the 1-persistence PFHFS scheme. Experimental results of the proposed framework are presented in section 3. Finally, the conclusion is depicted in section 4.

2. Methodology

In this paper, the proposed framework for TSA is depicted based on triple steps as shown in Fig. 1. In the first step, we define the reactive power-based feature and then generate the transient database via offline contingency simulation. Second, we offer the 1-persistence parallel fragmented hybrid feature selection scheme (PFHFS) for extracting discriminative features, and third, the efficacy of optimal feature set on TSA is evaluated through the cross-validation procedure.

2.1 Extraction of reactive power-based feature

Definition of the proper feature of power system components which is closely associated with transient analysis is considered as a preliminary step on TSA. For example, rotor angle and voltage magnitude are applied as the fixed features in most transient researches. However, the definition of the novel feature still remains an open issue in this scope. In this paper, we offer the reactive power-based feature according to the following basic reason.

In addition to the resistors elements in the power grid that always consume active power, other

elements such as inductor and capacitor have the amount of power received and power consumed in an equal manner. In fact, these elements caused a kind of exchange in the power grid when voltage is applied, namely, gains power when consumed and return power during production. This is known as reactive power. Such power exchange in transmission lines affects the power network in two ways negatively: takes up the transmission line capacity and waste power (The beer analogy). On the other hand, all of the required amounts of reactive power are not generated by power plants. In fact, if a capacitor is situated next to the inductor, the upstream electric power system could be free from reactive power. Consequently, other sources play a role in generating reactive power. Now we have to answer the question of what quota of reactive power consumed by loads is produced by power plants? Measuring based on this question can reflect the active power losses in the power system where network instability is closely related to active power. According to what was mentioned, the reactive power-based feature is shown as follow:

1) First univariate feature (Basic phrase): The proportion of total $QLOAD$ to total $QELEC$ which is defined as Eq. (1):

$$F_1^{tp} = \frac{\sum_{i=1}^{N_{busgen}} QLOAD_i}{\sum_{i=1}^{N_{busgen}} QELEC_i}; \quad p = 1.2 \dots r \quad (1)$$

Where $QELEC$ is the machine reactive power and

QLOAD is the reactive power consumption. Also, N_{busgen} is the number of bus generators in the test case and r is the data object in total length of feature in second (t_p).

2) Second univariate feature (Gradient phrase): The first-order derivative of the proportion of total *QLOAD* to total *QLEEC* which is defined as Eq. (2):

$$F_2^{t_p} = \text{Gradient of } F_1 = F_1^{t_{p+1}} - F_1^{t_p}; p = 1.2 \dots r \quad (2)$$

2.2 1-persistence PFHFS scheme

In this paper, the hybrid strategy including filter and wrapper techniques was considered for selecting MRFs of RP2vTS (See Fig. 2). Each univariate time series is individually entered into the optimal feature selection process in a parallel manner. As can be seen in Fig. 2, the 1-persistence PFHFS scheme consists of the conjoined steps as follow:

Step 1) Data fragmentation on $[data]_{(1122) \times 2 \times 50}$ (segmentation on instances and univariate time series):

$${}^j\text{Segment}_{[m \times z]}^i = \text{Fragmentation on } \langle [data]_{\text{sample}(1122) \times \text{univariate}(2) \times \text{feature points}(50)} \rangle \quad (3)$$

Where j indicates the number of univariate time series constrained $\{j | j = 1, 2\}$, i the segment number in j constrained $\{i | i = 1.2 \dots n\}$, m number of samples in ${}^j\text{Segment}^i$, and z number of feature points (in j univariate) selected for ${}^j\text{Segment}^i$. In fact, the dimensions per ${}^j\text{Segment}^i$ is $m \times z$.

Step 2) Applying information theory-based scheme (filter phase) per ${}^j\text{Segment}_{[m \times z]}^i$ to identify optimal features subset (OFSS) via cross-validation scenario:

a) First, we applied k -fold cross validation on ${}^j\text{Segment}_{[m \times z]}^i$:

$${}^j\text{fold}_{\text{train}: [m' \times z] - \text{test} [m'' \times z]}^i = k - \text{fold cross validation } [{}^j\text{Segment}_{[m \times z]}^i]; \quad (4)$$

Where k represents the number of desired partition or iteration on ${}^j\text{Segment}_{[m \times z]}^i$. These partitions called ${}^j\text{fold}_{\text{train}: [m' \times z] - \text{test} [m'' \times z]}^i$ where $v = 1:k \cdot m'$ and m'' number of samples for training and testing procedure in v iteration, respectively which is used for filter and wrapper phases.

b) Next, filter phase for each ${}^j\text{fold}_{\text{train}: [m' \times z]}^i$ related to ${}^j\text{Segment}_{[m \times z]}^i$ was applied:

$${}^j[\text{OFSS}\{1\# \}, \text{OFSS}\{2\# \}, \dots, \text{OFSS}\{h\# \}]^i = \text{filter phase } [{}^j\text{fold}_{\text{train}: [m' \times z]}^i]; h\#[1:s] \quad (5)$$

Where $h\#$ is the length of OFSS and s the maximum length of OFSS for each ${}^j\text{fold}_{\text{train}: [m' \times z]}^i$ per j .

Step 3) Applying wrapper phase (SVM classifier) on ${}^j\text{fold}^i$ for performance evaluation based on obtained $\text{OFSS}\{h\# \}$. First, we used Eq. (6) for training procedure. Next, testing procedure is conducted according to Eq. (7):

$${}^j\text{separating hyperplane}_{[m' \times \text{OFSS}\{1\# \}: \text{OFSS}\{h\# \}]}^i = \text{trainSVM } [{}^j\text{fold}_{\text{train}: [m' \times \text{OFSS}\{1\# \}: \text{OFSS}\{h\# \}]}^i] \quad (6)$$

$${}^j[\text{Acc}(\text{OFSS}\{1\# \}), \dots, \text{Acc}(\text{OFSS}\{h\# \})]^i = {}^j\text{separating hyperplane}_{[m' \times \text{OFSS}\{1\# \}: \text{OFSS}\{h\# \}]}^i [y] \quad (7)$$

Where $y = ({}^j\text{fold}_{\text{test}: [m'' \times \text{OFSS}\{1\# \}: \text{OFSS}\{h\# \}]}^i)$.

Step 4) Calculating mean value of OFSS performance with equal length situated in folds of ${}^j\text{Segment}_{[m \times z]}^i$, then return the maximum value of ${}^j\text{Mean}^i$:

$${}^j\text{Optimal Subset}_{h\#}^i = \max_{h\#[1:s]} \left(\begin{array}{l} {}^j\text{Mean}_{1\#}^i = \frac{1}{k} \sum_{v=1}^k {}^j[\text{Acc}(\text{OFSS}_{\{1\# \}})]^i \\ {}^j\text{Mean}_{2\#}^i = \frac{1}{k} \sum_{v=1}^k {}^j[\text{Acc}(\text{OFSS}_{\{2\# \}})]^i \\ \vdots \\ {}^j\text{Mean}_{h\#}^i = \frac{1}{k} \sum_{v=1}^k {}^j[\text{Acc}(\text{OFSS}_{\{h\# \}})]^i \end{array} \right) \quad (8)$$

Step 5) After extracting ${}^j\text{Optimal Subset}_{h\#}^i$ per j , we ranked it to find $l - \text{best}^1$ (F_1 univariate) and $l - \text{best}^2$ (F_2 univariate) as the Eq. (9) and Eq. (10):

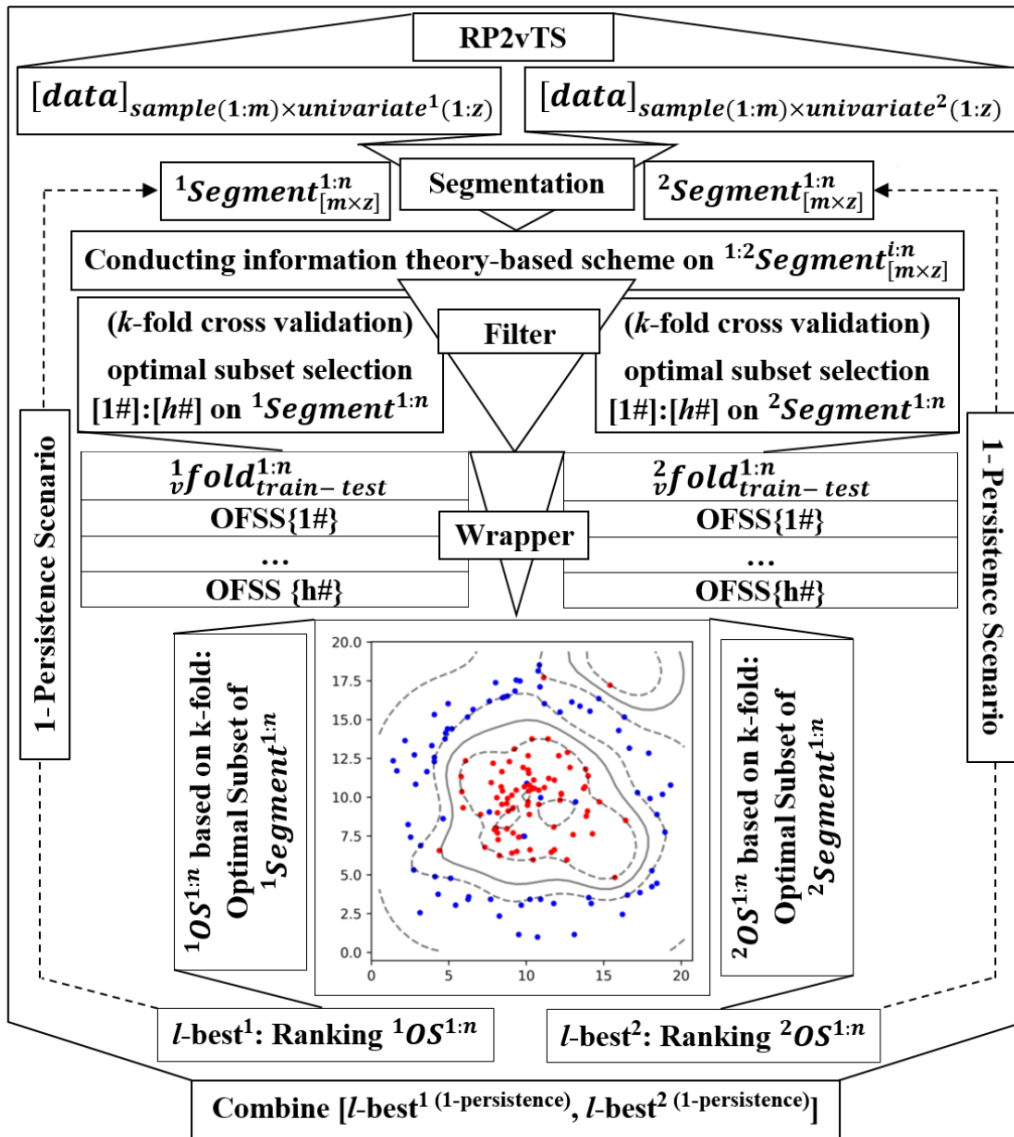


Figure. 2 Overall process for 1-persistence PFHFS scheme

${}^1Optimal\ Subset =$

$$Ranking_{l-best^1} \begin{pmatrix} \overbrace{\quad\quad\quad}^{h\#} \\ \overbrace{{}^1Optimal\ Subset^1}^{h\#} \\ \overbrace{{}^1Optimal\ Subset^2}^{h\#} \\ \vdots \\ \overbrace{\quad\quad\quad}^{h\#} \\ \overbrace{{}^1Optimal\ Subset^n}^{h\#} \end{pmatrix} \tag{9}$$

${}^2Optimal\ Subset$

$$= Ranking_{l-best^2} \begin{pmatrix} \overbrace{\quad\quad\quad}^{h\#} \\ \overbrace{{}^2Optimal\ Subset^1}^{h\#} \\ \overbrace{{}^2Optimal\ Subset^2}^{h\#} \\ \vdots \\ \overbrace{\quad\quad\quad}^{h\#} \\ \overbrace{{}^2Optimal\ Subset^n}^{h\#} \end{pmatrix} \tag{10}$$

Where l is the first l $\overbrace{\quad\quad\quad}^{h\#} Optimal\ Subset^j$ per j which is ranked by accuracy index.

Step 6) We apply obtained ${}^jOptimal\ Subset$ on PFHFS steps once again (1-persistence scenario) to achieve the final optimal feature subset per j .

After conducting the 1-persistence scenario, we combine $l - best^{1:2(1-persistence)}$ as the final optimal features subset. Then, according to Step 3 of Fig. 1, the final optimal features are used for transient stability prediction.

2.2.1 Filter phase: information theory

In the filter phase of 1-persistence PFHFS, statistical characteristics of features play a role in determining their importance. In fact, we focus on how attributes relate to the class target. To this end, relevance, interdependence, and redundancy analysis based on information theory criteria are considered. In terms of relevance, symmetric uncertainty (SU) [21] based on mutual information (MI) and entropy is used to measure the amount of information shared by two variables:

$$SU_{i,c}(F_i, C) = 2 \frac{MI(F_i, C)}{H(F_i) + H(C)} \quad (11)$$

Where F_i is feature and C target class.

Where $MI(F_i, C)$ and $H(C)$ is defined as Eq. (12) and Eq. (13), respectively:

$$MI(X, Y) = H(X) - H(X|Y) \quad (12)$$

Let X be a discrete random variable and probability density function $p(x) = \Pr\{X = x\}$:

$$H(X) = - \sum_{x \in X} p(x) \log p(x) \quad (13)$$

In terms of interdependence and redundancy analysis (IR), redundancy and the interdependent ratio between two features are used, which is given as Eq. (14):

$$IR(i, j) = 2 \frac{MI(F_i; C|F_j) - MI(F_i, C)}{H(F_i) + H(C)} \quad (14)$$

Where $-1 \leq IR(i, j) \leq 1$.

According to what was mentioned above, after the data fragmentation process, in the first step of the filter phase, the weight of features computed using SU measure (See Eqs. (11) to (13)) represents the degree of importance of the features. In the second step, as a supplementary analysis, redundancy and interdependence ratio has an influence on the weight of the features. First, the feature with higher SU is considered as an initial optimal feature ($OFFS \{1\# \}$). Then, for the rest of the features, the weight update

factor (WUF) is defined such that its amount is obtained from the redundancy and interdependence ratio (Eq. (14)) of each feature in the presence of the initial final feature. WUF is applied as a coefficient in SU of features, and the weight of features is updated. After sorting, the second final feature ($OFFS \{2\# \}$) is selected. This process will be continued until ($OFFS \{h\# \}$) features are selected. In this way, by applying the two-steps filter phase, s most important features among the initial features are selected and entered into the wrapper phase as input.

2.2.2 Wrapper phase: SVM classifier

One of the important approaches to cover the weakness of filter methods is applying learning models for selecting the optimal feature set. In fact, the performance evaluation of classification techniques in the presence of the selected features subset is considered a complementary step to explore optimal feature sets. Hence, after selecting the features subset in the filter phase, their effectiveness in increasing the accuracy of the classification technique in the training and testing procedure is examined. In this paper, the SVM classifier is used in the wrapper phase of the 1-persistence PFHFS scheme. The SVM is a supervised learning model based on Vapnik–Chervonenkis (VC) theory [22] for binary classification. Due to SVM's high-performance prediction on various scientific research, applying SVM for transient analysis is an incentive approach that has been considered by scholars on the TSA problem. On the other hand, in high-dimensional transient space, non-linear separability is the potential way to detect stability or instability accurately. To this end, kernel machine-based SVM to decrease the generalization error (maximum-margin hyperplane) is formulated following:

The optimization problem and the constraints of SVM is defined according to Eq. (15):

$$\alpha^* = \arg \min_{\alpha} \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \sum_{k=1}^l \alpha_k; \quad 0 \leq \alpha_i \leq C. \quad \sum_{j=1}^l \alpha_j y_j = 0. \quad i, j = 1, \dots, l \quad (15)$$

Where $K(x_i, x_j)$ is Gaussian radial basis function (RBF) kernel which is defined as Eq. (16):

$$K(x . x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right) \quad (16)$$

Where $\|x - x'\|^2$ is squared Euclidean distance between the two trajectory features. The optimal

$$f(x) = \operatorname{sgn}\left(\sum_{i \in S} \alpha_i y_i K(x_i . x) + b\right);$$

$$b = \frac{1}{S} \sum_{i \in S} \left[y_i - \sum_j \alpha_j y_j K(x_j . x_i) \right] \quad (17)$$

3. Experimental results

3.1 Transient dataset generation

The required transient dataset for the 1-persistence PFHFS scheme (See Fig. 1; step 2) and performance evaluation of the proposed framework (See Fig. 1; step 3) was described in this section. In this paper, transient simulations are conducted on the 16-Bus, 68-Machine, 5-Area system, a reduced-order equivalent of the New England-New York interconnection (NETS-NYPS) system. The single-line diagram (SLD) of the NETS-NYPS grid case is shown in Fig. 3 [23].

According to Fig. 4, the transient data generation procedure includes three steps: 1) run load flow, 2) sample transient disturbance characterized by its type, duration, and convert load options, and 3) save output channel related to QELEC and QLOAD. These steps are followed based on SIEMENS PSS/E application program interfaces (APIs) and Python technology. The APIs used in Python technology to automate dynamic simulation in PSS/E can be seen around the triple steps (e.g.; tysl, conl). All APIs based on the ‘psspy’ module for dynamic simulation using python scripting are shown in Table 1. We commented concise description per APIs (its function and argument) in Table 1 (# comments) according to Reference [24]. Consequently, 1122 simulation cases (578 case is stable and 544 case is unstable) was obtained based on defined RP2vTS for training and testing procedure. As can be seen in Table 1, the simulation time step is 0.0167 second (s), fault duration time almost 0.23. Also, we observed 50 cycles after fault clearing time (run simulation to 1.1027 s). Also, failure on any generator and line has been simulated on the N-1 contingency criterion. Furthermore, parameter tuning based on convert load

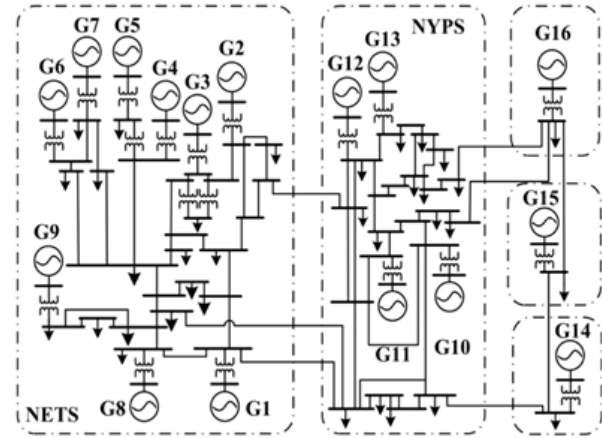


Figure. 3 SLD of NETS-NYPS grid case

for applying severe transient was considered in the form of four elements of “conl” activity namely convert the constant MVA load to constant current active/reactive load and constant admittance active/reactive load. According to what was mentioned above, for example, Fig. 5 shows the proportion of total QLOAD to the total QELEC (F_1 univariate of RP2vTS) for stable and unstable samples on NETS-NYPS grid case.

3.2 Extracting MRFs of RP2vTS by 1-persistence PFHFS

In this section, the result of applying the 1-persistence PFHFS scheme on RP2vTS is shown in Table 2. According to 1-persistence PFHFS scheme steps (See Section 2.2), the number of segments and size of its, maximum selected features by filter phase (s), type of cross-validation related to wrapper phase, and selected features based on filter and wrapper phase regarding PFHFS and 1-persistence PFHFS are shown in Table 2. After selecting the best subsets per time series features (F_1 & F_2) in PFHFS, these subsets are considered as input for the 1-persistence scenario. Finally, after conducting the 1-persistence scenario, we integrated selected best subsets per time-series features. In fact, F_1 cycle (22,26,27,28,42) and F_2 cycle (31,37) are final optimal cycles based on proposed feature selection scheme for transient stability prediction.

3.3. TSP results based on MRFs of RP2vTS

After extracting MRFs of RP2vTS via the 1-persistence PFHFS scheme, we evaluated the performance of the proposed framework for TSP. To this end, 5-fold cross-validation based on the SVM-RBF classifier was considered in this section. The prediction performance based on triple matrices (See

Table 1. Python scripting for dynamic simulation based on PSS/E API ('psspy' module)

```

for k in xrange(1,buscon+1):
    x=busm[0,k-1]
    psspy.psseinit(150000000) #Initialize PSS/E, requested bus size
    psspy.case(CASE) #PSS/E Saved Case file and transfers its data into the PSS/E working case
    report = r"C:\Users\...\report"+`k`+".txt"
    ierr=psspy.progress_output(2,report,[0,0]) #specify the progress output device; direct output to a file and
    open with carriage control format and, for files, for overwrite of existing files
    psspy.conl(-1,1,1) #initialize for load conversion
    psspy.conl(-1,1,2,[0,0],[100,0,0,100]) #convert the constant MVA load for a specified grouping of
    network loads to a specified mixture of the constant MVA, constant current, and constant admittance load
    characteristics
        psspy.conl(-1,1,3) # postprocessing housekeeping
    psspy.cong() #convert generators from their power flow representation in preparation for switching
    studies and dynamic simulations
    psspy.ordr() #calculate a sparsity preserving ordering of buses in preparation for the processing of
    network matrices
    psspy.fact() #factorize the network admittance matrix in preparation for switching studies and dynamic
    simulations
    psspy.tysl() #run switching study network solutions
    case_root=os.path.splitext(CASE)[0]
    psspy.save(case_root+"_C"+`k`+".sav") #save the PSS/E working case in a Saved Case File
    psspy.dyre_new(dyrefile="C:\Program Files (x86)\PTI\PSSE33\EXAMPLE\NETS-NYPS 68 Bus
    Syste.dyr") #read a dynamics data file, and place the model references specified on its data records into
    dynamics working memory
    psspy.chsb(sid=0,all=1,status=[-1,-1,-1,1,3,0]) #chsb: specify the simulation variables to monitor during
    dynamic simulation. run (activity chsb); machine reactive power: QELEC
    psspy.chsb(sid=0,all=1,status=[-1,-1,-1,1,26,0]) #chsb: specify the simulation variables to monitor during
    dynamic simulation. run (activity chsb); reactive power consumption: Qload
    psspy.snap(sfile="C:\Program Files (x86)\PTI\PSSE33\EXAMPLE\python_test_"+`k`+".snp") #save
    PSS/E dynamics working memory into a Snapshot File
    psspy.dynamics_solution_param_2(intgar=[11,271,0,0,0,0,1],realar=
    [1,0.0005,0.0167]) #modify the dynamic simulation solution parameters in dynamics working memory, e.g.;
    intgar[1] for network solution maximum number of iterations and intgar[2] for number of output channels
    being monitored
    psspy.strt(outfile="C:\Users\...\out\python_test_"+`k`+".out") #recorded during the dynamic simulation
    psspy.throwPsseExceptions = True #throw exceptions instead of returning error codes values
    psspy.run(tpause=0.03677) #fault occurrence time
    psspy.dist_bus_fault(ibus=x,units=3,values=[0,0]) #apply a fault at a bus during dynamic sim; ibus: #bus
    which the fault is to be placed; basekv: base voltage in kV used to calculate the per unit fault admittance.
    psspy.run(tpause=0.2677) #fault clearing time
    psspy.dist_clear_fault() #clear a fault during dynamic simulation
    psspy.dist_bus_trip(ibus=x) #disconnect a bus during dynamic simulation
    psspy.run(tpause=1.1027) #simulation time termination

```

Table 3) per splits (5-fold training and testing sets) was calculated via fine-tuning on parameters for the classifier (C and σ). The optimal pairs in SVM-RBF are selected from $\{2^i | i = 0, 1, 2, \dots, 8\}$ and $\{2^j | j = -5, -3, \dots, 4\}$, respectively. Fig. 6 shows the SVM-RBF performance for TSP based on fine-tuning (C , σ) grouped by folds. The maximum value of the Acc index (best values of C and σ) for each fold is obtained according to Table 4.

Also, TPR and TNR are measured according to obtained Acc per folds. The mean of indices (Acc , TPR , TNR) obtained in folds depicted in the

underline-face. As can be seen in Table 4, the classification accuracy of SVM-RBF based on MRFs shows the high-performance capacity for TSP (Acc : 96%, TPR : 95%, and TNR : 97%).

In addition to the importance of prediction accuracy on TSP, the processing time is also of great concern on real-time TSA. As stated in the third paragraph of the paper's introduction (applying feature selection scheme), the timely corrective control action (<1 s) [2] is possible through low processing time. As mentioned earlier, we observed 50 cycles after fault

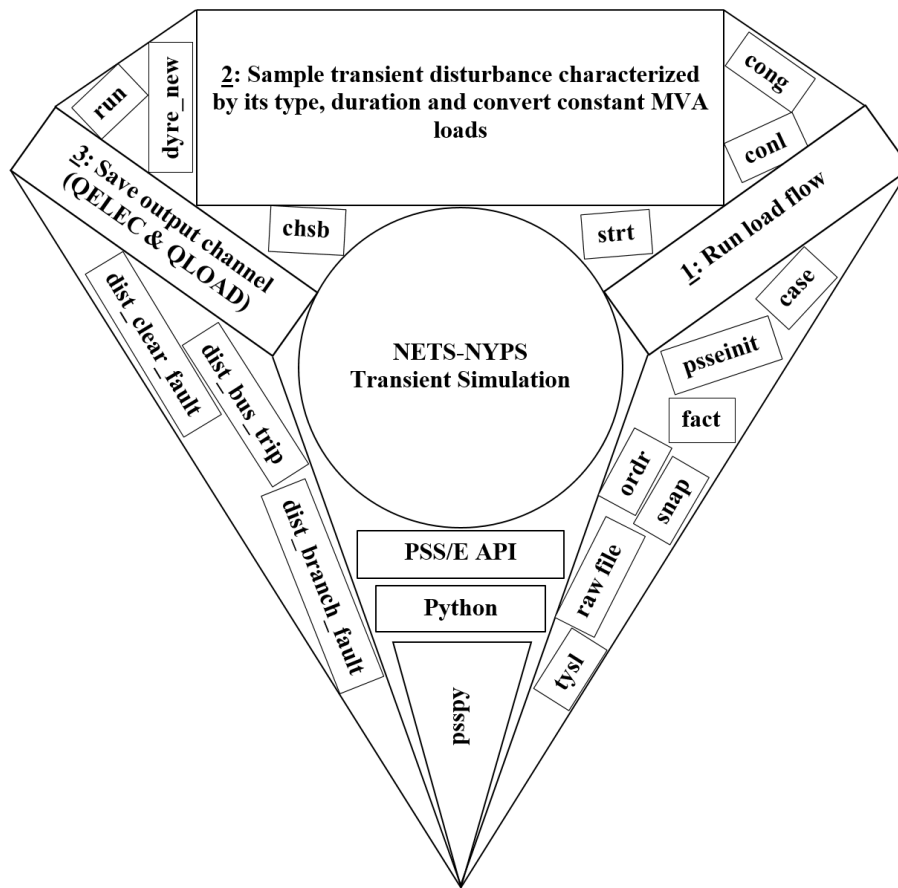


Figure. 4 Overall transient simulation on NETS-NYPS grid case

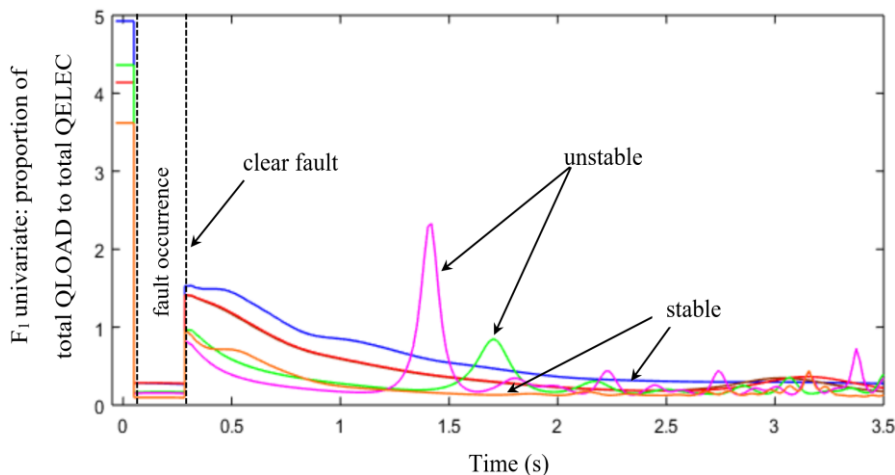


Figure. 5 F1 univariate fluctuations of RP2vTS: stable and unstable sample

clearing time. The transient space search on 50 cycles in the F1 time series and 50 cycles in the F2 time series for selecting MRFs, does not exceed 42 cycles (cycle 42 and less than 42 selected as the MRFs in the F1 time series, in the F2 time series cycles less than 42 are selected) (See Table 2). Hence, it is enough to observe 42 cycles after fault clearing time, which is about 701 milliseconds (ms) (See Table 5). Also, according to our measurements, the prediction time based on SVM-RBF regarding selected MRFs is 13.4

ms. Consequently, the processing time for TSP is 714.8 ms, which provides ample time to take corrective action.

3. 4. Comparison of experimental methods: 1-persistence PFHFS vs. three efficient FSS

Comparing the performance of 1-persistence PFHFS with three efficient FSS for TSA is

Table 2. The obtained MRFs based on 1-persistence PFHFS

F₁ time series: Max of Acc. based on SVM-cross validation on segments (fine-tuning of SVM-RBF parameters)				
	Segment	Size	Features (cycle)	[Selected features based on filter and wrapper phase] (Acc.)
PFHFS (F ₁ time series) (5-cross validation) (s: 4#)	Segment 1	115×5	[1,2,3,4,5]	[1,2] (89.56)
	Segment 2	115×5	[6,7,8,9,10]	[6,7,8,9] (88.69)
	Segment 3	115×5	[11,12,13,14,15]	[11] (86.95)
	Segment 4	115×5	[16,17,18,19,20]	[16,17] (98.26)
	Segment 5	115×5	[21,22,23,24,25]	[21,22,23] (100)
	Segment 6	115×5	[26,27,28,29,30]	[26,27,28,29] (100)
	Segment 7	115×5	[31,32,33,34,35]	[31] (100)
	Segment 8	115×5	[36,37,38,39,40]	[36] (97.39)
	Segment 9	115×5	[41,42,43,44,45]	[41,42] (97.39)
	Segment 10	87×5	[46,47,48,49,50]	[46,47,48,49] (90.71)
1-persistence PFHFS (F ₁ time series) (3-cross validation) (s: 2#)	Segment 1	270×3	[26,22,41]	[26,22] (90.37)
	Segment 2	270×3	[27,42,16]	[27,42] (98.51)
	Segment 3	270×3	[28,17,21]	[28] (96.66)
	Segment 4	312×3	[31,23,29]	[31,23] (90.06)
F₂ time series: Max of Acc. based on SVM-cross validation on segments (fine-tuning of SVM-RBF parameters)				
PFHFS (F ₂ time series) (5-cross validation) (s: 4#)	Segment 1	115×5	[1,2,3,4,5]	[1,2,3,4] (75.65)
	Segment 2	115×5	[6,7,8,9,10]	[6,7,8] (78.26)
	Segment 3	115×5	[11,12,13,14,15]	[11,12,13,14] (75.65)
	Segment 4	115×5	[16,17,18,19,20]	[16,17,18] (78.26)
	Segment 5	115×5	[21,22,23,24,25]	[21,22] (81.73)
	Segment 6	115×5	[26,27,28,29,30]	[26,27,28,29] (73.04)
	Segment 7	115×5	[31,32,33,34,35]	[31,32,33,34] (87.82)
	Segment 8	115×5	[36,37,38,39,40]	[36,37,38,39] (84.34)
	Segment 9	115×5	[41,42,43,44,45]	[41,42] (71.30)
	Segment 10	87×5	[46,47,48,49,50]	[46,47,48] (64.31)
1-persistence PFHFS (F ₂ time series) (3-cross validation) (s: 2#)	Segment 1	270×4	[6,16,21,36]	[6,16] (71.11)
	Segment 2	270×4	[17,7,31,37]	[31,37] (82.96)
	Segment 3	270×4	[8,32,38,64]	[32,38] (78.14)
	Segment 4	312×3	[18,33,34,39]	[18, 33] (66.98)
Integration of best [F₁ 1-persistence & F₂ 1-persistence]: (F₁ cycle: [22,26,27,28,42] + F₂ cycle [31,37])				

Table 3. The performance metrics

Metrics	Descriptions
Accuracy	Acc=(TP+TN)/(TP+TN+FP+FN)
Sensitivity	TPR=TP/(TP+FN)
Specificity	TNR=TN/(TN+FP)

Symbols; P: stable sample, N: unstable sample, T: predicted correctly, F: predicted incorrectly

considered in this section. Hence, we concentrated on the minimum redundancy maximum relevance (mRMR), ReliefF, and fast correlation-based filter (FCBF) algorithms which considered by scholars in the field of transient stability [17-20]. In the first step, three efficient FSS (3eFSS) was applied to RP2vTS for selecting MRFs.

Table 4. Results of transient stability status prediction based on MRFs of RP2vTS SVM-RBF (5-fold cross validation)

Max Acc. per fold: (Accuracy) [TPR/TNR]	
fold 1	fold 2
(96.22) [96.30/ 96.15]	(97.16) [96.30/ 98.08]
fold 3	fold 4
(97.16) [98.15/ 96.15]	(92.45) [88.89/ 96.15]
fold 5	
(96.22) [92.59/ 100]	
Mean (measures): Accuracy/TPR/TNR	
95.84/ 94.44/ 97.30	

Next, the obtained MRFs by 3eFSS entered to SVM-RBF classifier for performance evaluation on TSP. After conducting FSS methods on RP2vTS, the

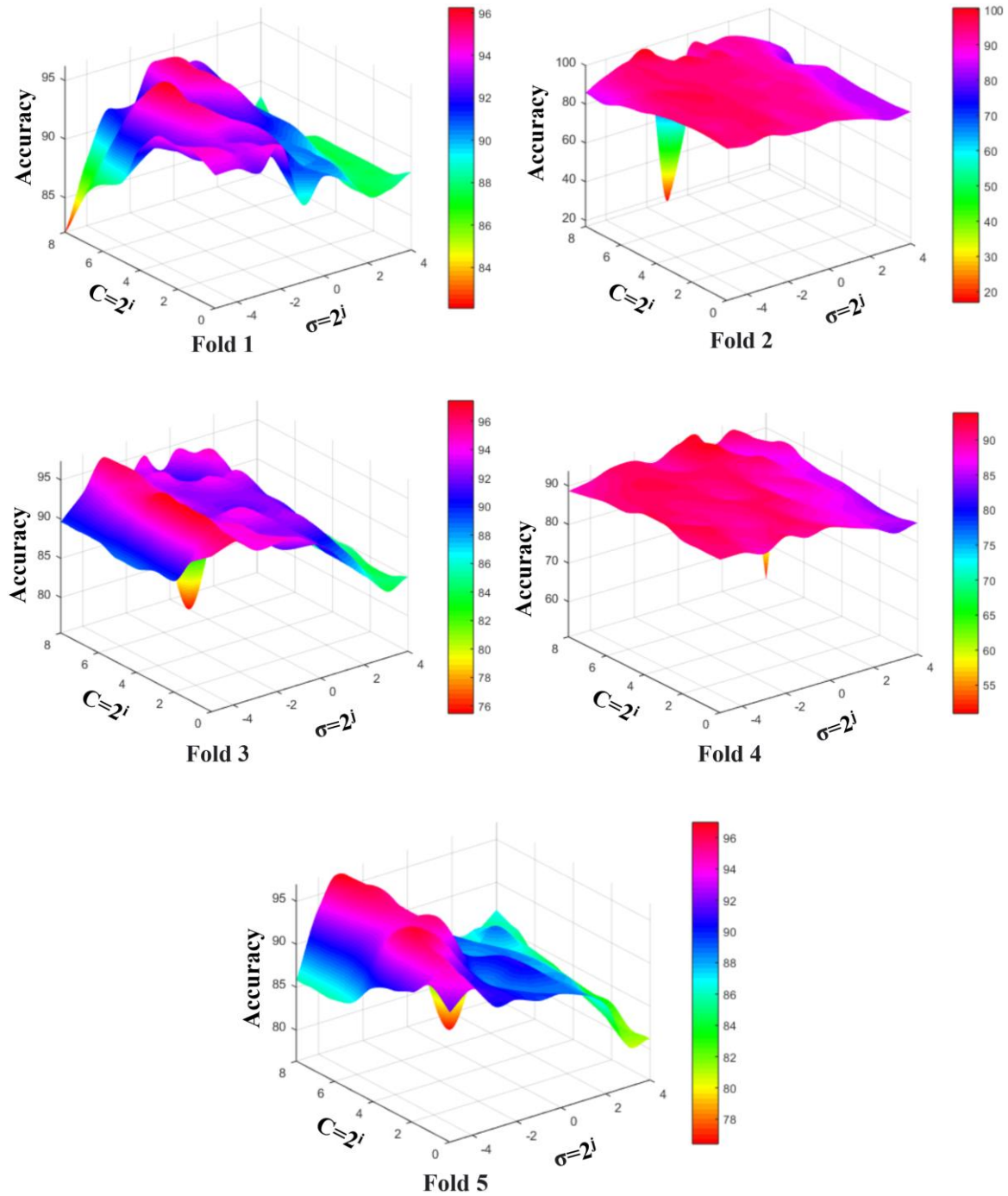


Figure. 6 Illustration of SVM-RBF performance for TSP based on learning parameters group by folds.

Table 5. Processing time for TSP based on MRFs

Observation in second / cycle	Processing Time ^a
42 / 0.7014	701.4 ms+13.4 ms= 714.8 ms

^aProcessing Time=PoT+prediction time; ^b millisecond

Table 6. Obtained MRFs grouped by FSS

Methods	Obtained MRFs : F1 cycle + F2 cycle
1-persistence PFHFS	F ₁ : [22,26,27,28,42]+F ₂ [31,37])
mRMR	F ₁ : [1,2,3]+F ₂ [5,6,7,31])
ReliefF	F ₁ : [2,3,4,5,7,9,10]+F ₂ [-])
FCBF	F ₁ : [6,8,16,17,18,19,20]+F ₂ [-])

MRFs grouped by 1-persistence PFHFS and 3eFSS was obtained according to Table 6. As can be seen in Table 7, 1-persistence PFHFS outperformed mRMR,

Table 7. Comparing 1-persistence PFHFS performance with mRMR, ReliefF, and FCBF.

Methods	5-fold cross validation				
	fold 1	fold 2	fold 3	fold 4	fold 5
1-persistence PFHFS	Max Acc. per fold: (Accuracy) [TPR/TNR]				
	(96.22)	(97.16)	(97.16)	(92.45)	(96.22)
	[96.30/ 96.15]	[96.30/ 98.08]	[98.15/ 96.15]	[88.89/ 96.15]	[92.59/ 100]
	Mean (measures): Accuracy/TPR/TNR				
	95.84/ 94.44/ 97.30				
mRMR	Max Acc. per fold: (Accuracy) [TPR/TNR]				
	(92.45)	(92.45)	(94.33)	(89.62)	(92.45)
	[92.59/ 92.31]	[94.44/ 90.38]	[98.15, 90.38]	[90.74, 88.46]	[88.89, 96.15]
	Mean (measures): Accuracy/TPR/TNR				
	92.26/ 92.96/ 91.53				
ReliefF	Max Acc. per fold: (Accuracy) [TPR/TNR]				
	(93.39)	(98.11)	(95.28)	(90.56)	(97.16)
	[98.15/ 88.46]	[98.15/ 98.08]	[96.30/ 94.23]	[88.89/ 92.31]	[94.44/ 100]
	Mean (measures): Accuracy/TPR/TNR				
	94.90/ 95.18/ 94.61				
FCBF	Max Acc. per fold: (Accuracy) [TPR/TNR]				
	(95.28)	(94.33)	(94.33)	(89.62)	(94.33)
	[94.44/ 96.15]	[96.30/ 92.31]	[94.44/ 94.23]	[87.04/ 92.31]	[90.74/ 98.08]
	Mean (measures): Accuracy/TPR/TNR				
	93.57/ 92.59/ 94.61				

ReliefF, and FCBF approaches based on triple metric (Acc, TPR, and TNR) regarding cross validation technique (ignoring only 0.74% less TPR than ReliefF). An important point to note is that among the 3eFSS applied cohesively on the whole of feature space, only the mRMR can extract the optimal cycles of the F_2 time series feature (See Table 6). In fact, based on ReliefF and FCBF methods, none of the cycles in F_2 survive. However, mRMR does not have a high performance compared to the FCBF and ReliefF algorithms (ignoring only 0.37% more TPR than FCBF) due to applying its own information theory formula on F_1 in shallow-mining manner. On the other hand, by sacrificing blurred MRFs of F_2 through F_1 -oriented filter mechanism in FCBF and ReliefF, the opportunity to gain high-performance capacity than 1-persistence PFHFS is lost. Overall, conducting the segment-based filter-wrapper scheme by 1-persistence PFHFS on dual time-series features separately caused to taking optimal cycles of F_1 and F_2 and achieving high accuracy than FCBF, mRMR, and ReliefF on TSP.

4. Conclusion

In this paper, the 1-persistence PFHFS scheme is offered for extracting optimal feature subset on RP2vTS for TSA in an accurate and timely manner. According to the proposed framework, first, we define reactive power-based two-variate time series (called RP2vTS). Next, the hybrid feature selection

scheme based on filter and wrapper methods is offered for selecting the most relevant features (MRFs) per univariate time series. The information theory concept was considered in the triple index in the filter phase, namely relevance, interdependence, and redundancy analysis. Next, the obtained preliminary optimal features subset entered into the wrapper phase for measuring classification performance based on selected subsets. Also, as the 1-persistence scenario, the obtained MRFs per univariate time series followed the PFHFS steps once again. Finally, after integrating the final optimal feature subset per univariate features, we evaluated the efficacy of the MRFs for TSP based on the SVM-RBF classifier. The results showed that the proposed framework by selecting MRFs of RP2vTS offers high-performance capacity (Acc: 96%, TPR: 95%, and TNR: 97%, processing time: 714.8 ms) on TSP. To ensure the efficacy of the 1-persistence PFHFS for TSA, experimental comparison in the presence of 3eFSS was considered as the final report in this paper. The results showed that the 1-persistence PFHFS outperformed mRMR, ReliefF, and FCBF algorithms based on triple metric regarding cross-validation technique.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, supervision, and project administration have been done by 1st author.

References

- [1] P. Kundur, J. Paserba, V. Ajjarapu, G. Andersson, A. Bose, C. Canizares, N. Hatziargyriou, D. Hill, A. Stankovic, C. Taylor, T. Van Cutsem, and V. Vittal, “Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions”, *IEEE Transactions on Power Systems*, Vol. 19, No. 3, pp. 1387–1401, 2004.
- [2] M. Pavella, D. Ernst, and D. Ruiz-Vega, *Transient Stability of Power Systems: A Unified Approach to Assessment and Control*, Springer, Boston, MA 2000.
- [3] Y. Zhou, J. Wu, L. Hao, L. Ji, and Z. Yu, “Transient Stability Prediction of Power Systems Using Post-Disturbance Rotor Angle Trajectory Cluster Features”, *Electric Power Component and Systems*, Vol. 44, No. 17, pp. 1879-1891, 2016.
- [4] F. R. Gomez, A. D. Rajapakse, U. D. Annakkage, and I. T. Fernando, “Support Vector Machine-Based Algorithm for Post-Fault Transient Stability Status Prediction Using Synchronized Measurements”, *IEEE Transactions on Power Systems*, Vol. 26, No. 3, pp. 1474-1483, 2011.
- [5] L. S. Moulin, A. P. A. D. Silva, M.A. El-Sharkawi, and R. J. Marks, “Support vector machines for transient stability analysis of large-scale power systems”, *IEEE Transactions on Power Systems*, Vol. 19, No. 2, pp. 818-825, 2004.
- [6] W. Hu, Z. Lu, S. Wu, W. Zhang, Y. Dong, R. Yu, and B. Liu, “Real-time transient stability assessment in power system based on improved SVM”, *Journal of Modern Power Systems and Clean Energy*, Vol. 7, No. 1, pp. 26-37, 2019.
- [7] S. K. Azman, Y. J. Isbeih, M. S. E. Moursi, and K. Elbassioni, “A Unified Online Deep Learning Prediction Model for Small Signal and Transient Stability”, *IEEE Transactions on Power Systems*, Vol. 35, No. 6, pp. 4585-4598, 2020.
- [8] L. Zhu, D. J. Hill, and C. Lu, “Hierarchical Deep Learning Machine for Power System Online Transient Stability Prediction”, *IEEE Transactions on Power Systems*, Vol. 35, No. 3, pp. 2399-2411, 2020.
- [9] J. C. Gonzalez-Torres, G. Damm, V. Costan, A. Benchaib, and F. Lamnabhi-Lagarrigue, “A Novel Distributed Supplementary Control of Multi-Terminal VSC-HVDC Grids for Rotor Angle Stability Enhancement of AC/DC Systems”, *IEEE Transactions on Power Systems*, Vol. 36, No. 1, pp. 623-634, 2021.
- [10] S. Dasgupta, M. Paramasivam, U. Vaidya, and V. Ajjarapu, “PMU-Based Model-Free Approach for Real-Time Rotor Angle Monitoring”, *IEEE Transactions on Power Systems*, Vol. 30, No. 5, pp. 2818-2819, 2015.
- [11] A. F. Diaz-Alzate, J. E. Candelo-Becerra, and J. F. Villa Sierra, “Transient Stability Prediction for Real-Time Operation by Monitoring the Relative Angle with Predefined Thresholds”, *Energies*, Vol. 12, No. 5, pp. 838, 2019.
- [12] J. Zhao, Y. Zhu, and J. Tang, “Transient Voltage and Transient Frequency Stability Emergency Coordinated Control Strategy for the Multi-infeed HVDC Power Grid”, In: *Proc. of IEEE Power & Energy Society General Meeting (PESGM)*, Montreal, QC, Canada, 2020.
- [13] T. Liu and X. Wang, “Transient Stability of Single-Loop Voltage-Magnitude Controlled Grid-Forming Converters”, *IEEE Transactions on Power Electronics*, Vol. 36, No. 6, pp. 6158-6162, 2021.
- [14] F. Milano and Á. O. Manjavacas, “Frequency-Dependent Model for Transient Stability Analysis”, *IEEE Transactions on Power Systems*, Vol. 34, No. 1, pp. 806-809, 2019.
- [15] A. Mitra and D. Chatterjee, “Active Power Control of DFIG-Based Wind Farm for Improvement of Transient Stability of Power Systems”, *IEEE Transactions on*

- Power Systems*, Vol. 31, No. 1, pp. 82-93, 2016.
- [16] A. Mujcinagic, M. Kusljugic, and E. Nukic, "Wind Inertial Response Based on the Center of Inertia Frequency of a Control Area", *Energies*, Vol. 13, No. 23, pp. 6177, 2020.
- [17] L. Ji, J. Wu, Y. Zhou, and L. Hao, "Using Trajectory Clusters to Define the Most Relevant Features for Transient Stability Prediction Based on Machine Learning Method", *Energies*, Vol. 9, No. 11, pp. 898, 2016.
- [18] A. Stief, J. R. Ottewill, and J. Baranowski, "Relief F-Based Feature Ranking and Feature Selection for Monitoring Induction Motors," In: *Proc. of International Conf. on Methods & Models in Automation & Robotics (MMAR)*, Miedzyzdroje, Poland, 2018.
- [19] X. Li, Z. Zheng, L. Wu, R. Li, J. Huang, X. Hu, and P. Guo, "A Stratified Method for Large-Scale Power System Transient Stability Assessment Based on Maximum Relevance Minimum Redundancy Arithmetic", *IEEE Access*, Vol. 7, pp. 61414-61432, 2019.
- [20] J. Yan, C. Li, and Y. Liu, "Deep learning based total transfer capability calculation model", In: *Proc. of International Conf. on Power System Technology (POWERCON)*, Guangzhou, China, 2018.
- [21] Y. Piao and K. H. Ryu, "A Hybrid Feature Selection Method Based on Symmetrical Uncertainty and Support Vector Machine for High-Dimensional Data Classification," In: *Proc. of Asian Conf. on Intelligent Information and Database Systems*, Kanazawa, Japan, pp. 721-727, 2017.
- [22] V. N. Vapnik and A. Ya. Chervonenkis, "On the Uniform Convergence of Relative Frequencies of Events to Their Probabilities", *Theory of Probability & Its Applications*, Vol. 16, No. 2, pp. 264-280, 1971.
- [23] *Inter-connected New England test system (NETS) and New York power system (NYPS)*, <https://electricgrids.engr.tamu.edu/electric-grid-test-cases/new-england-68-bus-test-system/>
- [24] Siemens Industry, Inc., Siemens Power Technologies International, *PSS/E 33.4 application program interface (API)*, www.siemens.com/power-technologies, Schenectady, NY, USA, 2013.