Shahid Rajaee
Teacher Training University

Journal of
Electrical and Computer
Engineering Innovations
(JECEI)

Vol. 8 No. 2 Summer-Fall 2020

Journal of Electrical and Computer Engineering Innovations Vol. 8, No. 2, Summer-Fall 2020

JECEI

Shahid Rajaee
Teacher Training University

*p*ISSN 2322-3952
*e*ISSN 2345-3044

Journal of
Electrical and Computer
Engineering Innovations (JECEI)

Semiannual Publication

Volume 8, Issue 2, Summer-Fall 2020

Journal of Electrical and Computer Engineering Innovations
Vol. 8; Issue 2: 2020

Table of Contents

Journal of Electrical and Computer Engineering Innovations
(JECEI)
Journal homepage: http://www.jecei.sru.ac.ir

Research paper

Depuration based Efficient Coverage Mechanism for Wireless Sensor Network

S. Ashraf[1,*], T. Ahmed[1], Z. Aslam[2], D. Muhammad[3], A. Yahya[4], M. Shuaeeb[4]

[1]*College of Internet of Things Engineering, Hohai University Changzhou, Jiangsu, China.*

[2]*Petroweld Kurdistan Erbil, Iraq.*

[3]*Pakistan Steel Mills Karachi Pakistan.*

[4]*Dow University of Health Sciences Karachi Pakistan.*

Article Info

Article History:
Received 11 October 2019
Reviewed 12 December 2019
Revised 17 February 2020
Accepted 18 April 2020

Keywords:
Coverage tribulations
Canny solution
Cost-effective
Sensor nodes
Coverage range
Iterations
Depuration

*Corresponding Author's Email Address:
nfc.iet@hotmail.com

Abstract

Background and Objectives: The quick response time and the coverage range are the crucial factors by which the quality service of a wireless sensor network can be acknowledged. In some cases, even networks possess sufficient available bandwidth but due to coverage tribulations, the customer satisfaction gets down suddenly. The increasing number of nodes directly is neither a canny solution to overcome the coverage problem nor a cost-effective. In fact, by changing the positions of the deployed node sagaciously can resolve the coverage issue and seems a cost-effective solution. Therefore, keeping all circumstances, a Depuration based Efficient Coverage Mechanism (DECM) has been developed. This algorithm suggests the new shifting positions for previously deployed sensor nodes to fill the coverage gap.

Methods: It is a redeployment process and accomplished in two rounds. The first round avails the Dissimilitude Enhancement Scheme (DES), which searches the node to be shifted at new positions. The second round controls the unnecessary movement of the sensor nodes by the Depuration mechanism thereby the distance between previous and new positions is reduced.

Results: The factors like loudness, pulse emission rate, maximum frequency, and sensing radius are meticulously explored during simulation rounds conducted by MATLAB. The performance of DECM has been compared with superlative algorithms i.e., Fruit Fly Optimization Algorithm (FOA), Particle Swarm Optimization (PSO), and Ant Colony Optimization (ACO) in terms of mean coverage range, computation time, standard deviation, and network energy diminution.

Conclusion: According to the simulation results, the DECM has achieved more than 98% coverage range, with a trivial computation time of nearly 0.016 seconds as compared to FOA, PSO, and ACO.

Introduction

Usually, wireless sensor networks (WSNs) are incorporating with small-sized self-governing wireless sensor device, which is generally placed in aggressive and vulnerable environments to monitor and collect the data. However, in spite of widespread adaptation, WSNs

are given to multiple restrictions associated with processing abilities, thin wireless bandwidths, random sensor node deployment, limited storage spaces, and limited battery power. The fundamental issue in observing such environments is the area coverage which reflects how well the region is monitored. Coverage is usually defined as a measure of how well and how long the sensors are able to observe the physical space [1].

The quality of coverage in static sensor is significantly affected by the initial deployment location of the sensors. Unfortunately, sensor deployment cannot be performed manually in most applications [2], for instance, the deployment in disaster areas, harsh environments, and toxic regions. Most of the previous studies showed that, sensors were usually deployed by scattering from an aircraft; however, the actual landing position cannot be uniform due to the existence of obstacles for instance, buildings, trees and wind causing some areas of the sensing region to be denser than others. Therefore, even if a large number of redundant nodes are deployed, the desired level of coverage still cannot be achieved. Therefore, it is essential to make use of sensors, which can move iteratively to a better location that can give the required coverage [3].

In order to address the sensing coverage area, it is important to understand the mobility control attribute [4], of the sensor nodes. Indeed, sensor nodes have two type of mobility control attributes i.e., centralized and distributed [5]. Regarding centralized attribute, the bunch of nodes are centrally monitored by a sink node that overhears the sensing data from neighbouring nodes while in distributed networks, the sensors are self-controlled.

All sensor nodes have limited sensing and communication abilities [6], which make the sensor nodes unable to obtain the entire network information. Due to which sensors are deployed randomly and allowed to move and communicate with their neighbours by exchanging information between them [7]. The miniaturized robotics have overcome some hurdles regarding sensors mobility. Thereby, mobile sensors have the same sensing capability as static sensors [8], and can move freely to correct locations for providing the required coverage. On the other hand, it is not a cost-effective solution.

Keeping all aforementioned challenges, it is motivated to design a sagacious sensor node deployment strategy which should enhance the coverage area by consuming just confine energy metrics. Considering the pattern of a hybrid sensor network [9], which composed of mobile and static sensors we have proposed a Depuration based Coverage Mechanism for Wireless Sensor Network (DECM). For this purpose, a DECM algorithm has been designed which focus how to redeploy the sensor nodes to improve network area coverage in hybrid WSNs environment. It is indeed a cost-effective solution towards improving coverage with unevenly deployed sensors.

Initially, algorithm aims to determine where the sensor nodes should be moved while incurring the trivial moving cost. This will result only a confine moving cost including the accumulated moving distance, total number of moves, and communication rounds. The proposed DECM mechanism ultimately can maintain a balance between coverage with confine resource consumption during node redeployment process.

A. Working mechanism of proposed DECM

Initially, the nodes are deployed with some random positions, with certain velocities [10], to search the shrewd target positions in network coverage area. The minimum distance value and related coordinates are being recorded. After getting best minimum distance value the intended positions are crosschecked otherwise process will be repeat the same step. The further proceedings are explained stage by stage through Flow chart shown as Fig. 1.

Stage 1: Initialize all the parameters including the group size (n), the maximum number of iterations and the initial positions of sensor node group ($Xinitial$, $Yinitial$), step length [11], number of area range points, loudness and pulse rate, minimum and maximum frequency, upper and lower bounds [12]. All these parameters are being calculated through (1), (2) ,where i varies from 1 to n, LB and UB is lower and upper bounds, and n is the size of sensor node group.

$$X_{initial}(i) = LB + (UB - LB) * Random\ value \quad (1)$$

$$Y_{initial}(i) = LB + (UB - LB) * Random\ value \quad (2)$$

Stage 2: The essential parameters of the sensor nodes like positions (xit), velocities (vti) and frequencies for time t are updated as expressed in (3)-(5),

$$f_i = f_{min} + (f_{max} - f_{min})\beta \quad (3)$$

$$V_i^t = V_i^{t-1} + (x_i^t - x^*)f_i \quad (4)$$

$$x_i^t = x_i^{t-1} + v_i^t \quad (5)$$

where β is an arbitrary vector whose value is lies between 0 and 1, the $fmax$ represents maximum frequency and $x*$ indicate the Shrewd solution.

Stage 3: The distance of all the sensor nodes from the current area position (Distn*m) [13], is being computed by the (6)

$$Dist_{n*m} = \sqrt{(X_{n*m} - x_j)^2 + (Y_{n*m} - y_j)^2} \quad (6)$$

where Xn*m and Yn*m are initial positions of n*m sensor nodes xj and yj are coordinates of j area range.

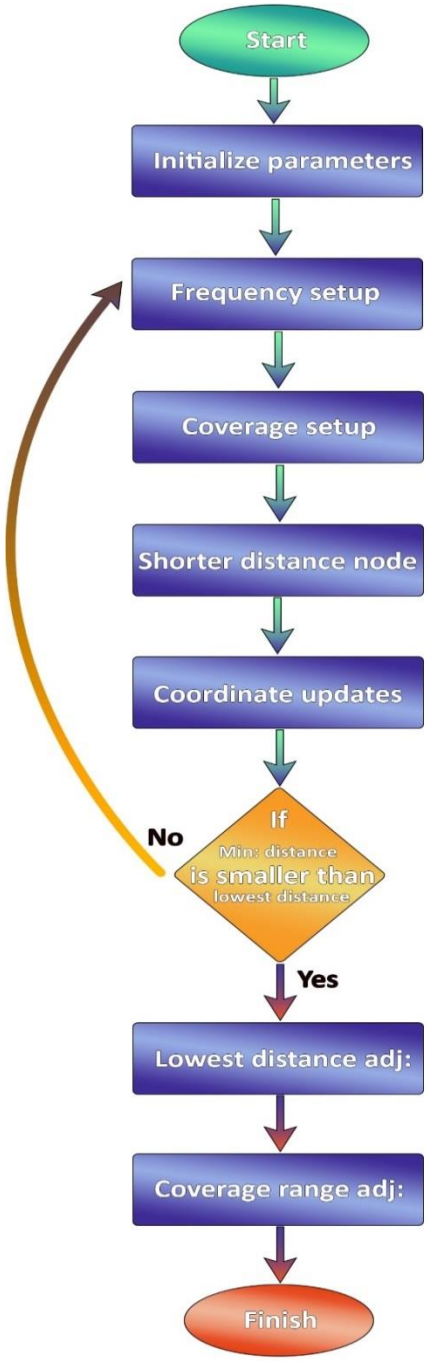


Fig. 1: DECM information flow chart.

Stage 4: Any sensor node having minimum distance value to the intended node positions are compared and this moving distance is selected.

Stage 5: The lowest distance value and related coordinates are recorded in corpus Table.

Stage 6: The lowest and shrewd distance value is compared with other distance value during every iteration. If no other shrewd distance value is found then this lowest and shrewd value and its coordinates are updated and sensor node shift its position to the intended target in accordance to the condition defined in DECM algorithm otherwise, repeats step from 2 to 5.

Stage 7: The overall network Coverage Range (CR) [14], has been computed through (7).

$$CR = \frac{N_{m*n}}{M*N} \quad (7)$$

The M*N is the network coverage area, m*n represents total summation points of each sensor node.

Further explanation is given in the third section. Our unique contributions have been summarized as:

• The proposed DECM algorithm tends to overcome related issues with the network coverage range by shifting already deployed sensor nodes from previous to new positions.

• In some cases, it makes substitution of nodes to adjust the coverage hole.

• The unnecessary sensor movement is also being monitor to reduce the movement distance between nodes which prevents the wastage of the energy resource.

• The simulation results generated through Matlab has vouched the succulent performance of DECM when compared with previous work FOA, PSO and ACO.

• The proposed DECM algorithm accomplished the operation in two junctures, during first juncture the intended target positions of the sensor node is computed through Dissimilitude Enhancement Scheme (DES). The second juncture is referred as Depuration, where the moving distance between node is sagaciously reduced, thereby the target positions are achieved.

The rest of the manuscript is structured as: The previous work has been rummaged out in the second section, the proposed methodology has been explained in the third section, while in the fourth section renders the output performance and the result discussion. Finally, overall achievements have been summarized in the form of conclusion in the next section.

Literature Review

Usually sensor nodes are deployed to cover the area between distinct boundaries. However, selection of most suitable area is ever remained a challenge [15]. In order to achieve the sufficient coverage area, the distributed deployment strategy is commonly used to improve the area coverage by moving the sensor nodes from one location to another. For this purpose, the distributed movement algorithms are being used wherein the coverage area is allocated in multiple segments. If any sensor node was unable to detect the event happenings within the deployed segment, no other sensor node can detect it. Eventually, the monitoring of each segment area for coverage gape (hole) and calculation of new location is the prime liability of the deployed sensor node. All distributed movement algorithms are facing numerous tribulations regarding new position calculation within the segment area while relocating the

new location. No researcher could ever address to overcome the nodes position reallocation challenge in hybrid environment. Therefore, no wireless network having coverage holes, can successfully carry out its monitoring operation [16]. The researcher tried to incorporate more iterations in their designed model to address the new allocation issue but it drastically increased the implications and causing higher energy consumption. To some extent, overcome these issues the numerous researchers have made substantial contributions. For example, the motion capability of sensor nodes with relocating ability and dealing with sensor failure have been identified by Qingguo et al. [17], They suggested a two-phase sensor relocation solution. The redundant sensors are first identified and then relocated to the target location. They proposed a grid-quorum solution to locate the closest redundant sensor, and proposed to use cascaded movement to relocate the redundant sensor. In fact, their suggested model could not control the exorbitant energy drainage and thereby whole network might die after few transmission rounds. On the other hand, Li Jun et al. [18], tried to address the coverage and load balancing issues by minimizing the moving distance and argued a centralized movement solution, based on the Hungarian method. However, the centralized movement technique revealed those sensor nodes having already appropriate positions when impelled to leave the position creating energy holes. Wang et al. [19], proposed three different distributed movement assisted sensor deployment algorithms, VEC, VOR, and Minimax, to improve the total area coverage. Thereby they used the Voronoi diagram to partition the monitoring area into n convex polygons where every polygon enclosed one sensor node only. This method utilizes the local polygon information to calculate the new position location to move sensor node. The VEC approach uses virtual force between two nodes to push them away from each other at a certain distance. Minimax and VOR algorithms are greedy, and try to fix the largest coverage hole by moving sensor node towards the farthest polygon vertex. The nodes approaching to the polygon do not need to move towards the farthest vertex.

As a result, this movement may not reduce coverage hole, but might increases the complications. The identification of new node location and its relative computation has been calculated through four local displacement conditions by the H. Mahboubi et al. [20], taking into account the circles having centered position within the respective polygons. Some centers might lie out of the polygon and thereby sensor nodes locating around those circles may not have movement. Consequently, this issue demands more rounds to overcome the coverage tribulation. The more the rounds it demands, the more the resources are being consumed; As a result, the sensor nodes will cause the network to confine the lifespan before the specified time. In order to increase the coverage rate of sensor nodes, various researchers have proposed different optimization techniques. A sensing and perception-based Fruit Fly Optimization Algorithm (FOA) was applied by Wen-Tsao Pan [21] , to address the position issue of the sensor node which aims to enhance the coverage matter in ideal and obstacle environment. As the fruit flies can reach the food source by using their smell and vision organs. Initially, they use osphresis organs to find all kinds of scents in the air. Then they fly toward to food. When they get close to the food, they use their vision organs to get closer. Similar action is adopted for relocating the sensors positions. Despite its advantages, there are critical issues for instance, the first pointing location remains poor. Further, the algorithm significantly traps into local optimum and the update strategy is limited.

An Edge Based Centroid (EBC) algorithm is proposed by Muhammad Sirajo et al. [22], and author claims about enhanced area coverage of monitoring field with minimal energy consumption.

In fact, this algorithm is based on Voronoi diagram that partitions the sensing field into polygons with one sensor node each to monitor any event in its respective subregion. The sensor node moves to new location at the center of each polygon from location, which improves area coverage. This algorithm depends on certain group of rules that ensures about the center of each polygon before the movement and thereupon the ratio of energy consumption can be lowered. Though this algorithm works smooth but no control over the uncouth movement of the node is addressed due to that sometime a node can make unusual and large displacement which might cause the energy wastage.

In pursuit of a better coverage technique, a majority of scholars have tried to use intelligent algorithms like, Genetic Algorithm (GA) [23], and Particle Swarm Optimization (PSO) [24], to solve the issue. Though, fruit fly algorithm is simple and practicable than GA and PSO but due to unavoidable limitations the researcher is still exerting their efforts to develop shrewder algorithm. Table 1, exhibits various comparison among such algorithms and shows a significant improvement by the proposed algorithm.

Table 1: Comparison of proposed DECM with in-practice algorithms

Algorithm	Working ground	Expediency	Impairments	Comparison with proposed DECM
Genetic Algorithm (GA)	Stochastic search methodology through generic system, withing a population it impels the recombination and mutation.	It is faster and have ability to find best quality solution in trivial time, possessed parallel capabilities. Easily discovers the global optimum.	Never guarantee for optimal solution. Hard to choose parameters like number of generations, population size. It is expensive.	It functions in hybrid environment, ensures about relocating the intended nodes position within the coverage area therefore energy consumption remains confined.
Particle Swarm Optimization (PSO)	Inspired by bird flocking and fish schooling. The particles move in a multidimensional search space and single intersection of all dimensions forms a particle.	It can overcome the unconstrained minimization issue. Provides the derivative free technique, it is less sensitivity, less dependent of a set of initial points. It can generate high-quality solutions.	It can easily fall into local optimum in high-dimensional space and has a low convergence rate in the iterative process. Difficult to adopt the best topology.	At the beginning it rummages where the sensor nodes should be moved therefore local minima can easily be avoided.
Bacterial Foraging Algorithm (BFA)	It works on search and optimal foraging decision making capabilities, problems, movement take place either in clockwise or counter clockwise direction	Used for unconstrained numerical optimization, having dual movement i.e., swimming and tumbling called chemotaxis,	Having weak ability to perceive the environment and vulnerable to perception of local extreme, hard to deal with complex optimization problems	As it operates in two stages, thereupon no vulnerabilities can slow down the performance, each stage performs independently.
Ant Colony Optimization (ACO)	Based on social behaviour of the insects, the optimization process is initialized by random solutions	Rapid discovery of good solutions with guaranteed convergence,	Dependent sequences of random decisions, having complicated theoretical analysis, uncertain time to convergence	The Depuration technique in second stage reduce the moving distance and there exists no uncertainty.
Artificial Bee Colony (ABC)	Search optimization consists of three essential components: employed and unemployed foraging bees, and food sources.	It minimizes the expense of deploying nodes inside the monitoring regiondeals with local , solution, havingbroad applicability, complex functions	Slow process, higher number of objective function evaluation, number of dimensions might change	It maintains the network dimension by reducing the moving distance between the nodes.
Jenga-inspired optimization algorithm (JOA)	Based on greedy fast convergence, select minimum cost node subset through the roulette method, bridge between optimal solution and short computation time.	Address the Energy-Efficient Coverage issues, having stochastic approach to conduct random exploration, if sensor node cannot cover an area the other node take avail the chance	The detection probability decreases exponentially as the distance becomes greater	Have shrewd control over moving distance therefore no uncouth movement can degrade the overall communication.

Coverage Model

A coverage model explains the possible coverage range by the sensor nodes in coverage area.

All sensor nodes have various coverage range characterized by area where these sensors are being deployed, the accuracy, the environment factors and resolution. The coverage area depends on various

factors such as the signal strength generated from the source, distance between the sensor node and source and the rate of attenuation in propagation. For example, an acoustic sensor network establishing the coverage range to detect the mobile vehicles, the sensor nearer to a vehicle can detect higher acoustic signal strength than the one farther away from the vehicle due to signal attenuation, and as a result there is higher confidence of detecting vehicles.

A. Problem Formulation

For proposed coverage model, a two-dimensional coverage area has been considered. Further, the coverage area is divided into various segments each having unit size. When n number of sensor nodes have been deployed in targeted area m, thereby a full couplet of sensor node can be defined as given in (8),

$$S = \{S_1, S_2, \ldots S_n\} \tag{8}$$

the position of ith node is defined as Si = (xi , yi) where i = (1,2,…n). The coverage range of sensor Si can be expressed as a circle centered at its coordinates (xi ,yi) with the radius of the sensing range Rs. Let Ei, being a random variable for an event that a sensor node Si covers an area of segment A(xA, yA). The Presage factor for event Ei can be written as P{Ei} which is equal to the coverage presage i.e., P(Si, xA, yA). Thereupon, the happening of an event presage can be defined by the discrete coverage model expressed in (9).

$$\mathrm{P}(Si, xA, yA) = \begin{cases} 1, d(Si, xA, yA) \leq R_s \\ 0, other\ case \end{cases} \tag{9}$$

The Euclidean distance [25], of i^{th} sensor node from segment area A(*x,y*) can be computed by (10).

$$\mathrm{P}(Si, xA, yA) = \sqrt{(x - x_i)^2 + (y - y_i)^2} \tag{10}$$

All coverage pints within the coverage range are measured as unity covered [26], by the particular sensor whereas, the points outside of this coverage range is regarded as 0. The shrewd objective of coverage optimization issue is to provide sufficient Coverage Range (*CR*), by using a smaller number of sensor nodes. The *CR* is used to estimate the performance of sensor network. Generally, it is assumed that segment area point can be covered by any sensor node only once.

B. The proposed DECM Model

At present, among all optimization algorithms the DES [27], is considered as a fasted optimization scheme therefore we found it sagacious and motivated to take full advantage for our proposed DECM algorithm. Thus, the coverage range tribulations in WSN is being resolved by redeployment of sensor nodes through DES strategies and therefore the stages of DECM design model are being explained one by one.

Stage 1. Locating intended target positions of the node:

The depuration based efficient coverage mechanism (DECM) is an investigative search technique that utilizes the shrewd coverage mechanism. It exploits the position of the sensor node for potential solutions, individuals, to probe the search range. It initializes the parameters while addressing the coverage area issue as depicted in (11),

$$X_s = (x_{s1}, \ldots, x_{s2}, \ldots, x_{s3}) \tag{11}$$

considering 1 ≤ s, as the area range and $x_{s2} \in [a_s , b_s]$, where "a_s" and "b_s" denotes the lower and upper bound of the s^{th} node, respectively. After every transmission round t, the corresponding re-allocation round presages the intended position of the bodacious node which is expressed as (12),

$$V_s\,(t + 1) = X_{bodacious} + F(X_{r2}\,(t) - X_{r3}\,(t)) + F(X_{r4}\,(t) - X_{r5}\,(t)) \tag{12}$$

The $X_{bodacious}$ indicates the appropriate position of the node while *r* represents the transmission round and *F* points a scaling factor that is a distance control parameter between initial and the new node position. To increase the sensing range, the position parameter $V_s(t + 1)$ is incorporate the value of predicted node $X_s(t)$, thereby yields a temporal position $Q_s(t + 1)$ as expressed in (13),

$$Q_{s,j}\,(t + 1) = \begin{cases} V_{s,j}(t + 1), if\,(rand[0,1] \\ < FCR\ orj \\ = J_{rand})X_s, j(t), for\ other\ case \end{cases} \tag{13}$$

The rand (0,1) represents a uniformly distributed random positions, while j_{rand} exhibits randomly predicted positions within the range [1,D]. The FCR came up as a Fractional Control Parameter ∈ [0, 1], which shows the inherited characters of previous node position.

Proceeding towards final position, the temporal position $Q_s(t + 1)$ is being compared with predicted node $X_s(t)$. The newly generated position that possessed greater fitness metric among rest of the positions is our intended position of the node given in (14),

here $f(X)$ represnets the intended target position of the node. In fact, the sensor network performs the virtual movement and as long it achieves the intended position of the sensor node in accordance to the (14), physical displacement has been performed accordingly.

$$X_s\,(t + 1) = \begin{cases} Q_s(t + 1), if\big(f(Q_s(t + 1)\big) \geq f(X_s(t))) \\ X_s(t), other\ case \end{cases} \tag{14}$$

Stage 2. Depuration process:

The depuration process is performed to reduce the moving distance of the node. This will reduce the number of sensor nodes that need to move, as well as

reduce the average moving distance; however, it does not affect the network coverage. The moving distance reduction strategy can be understood as: Consider the initial positions of the deployed sensor nodes illustrated in Fig. 2. Senosr node s_1 is lying at position-1, s_2 with position-2, s_3 with position-3, s_4 with position-4, and s_5 with position-5. The sensor node s_1 is trying to move at new intended position i.e., intended-SP1. At the same time, another sensor node s_2 also trying to capture the same position but DECM algorithm systematically controls the movement of sensor node that are needed to be moved. The sensing range may even be fully overlapped by other nodes, these nodes are called redundant nodes. If coverage range $R_{area}(S)$ presages no substantial change to position of a sensor node is required when a node with smaller distance has already accessed the intended position thereby node s_2 can be removed from the queue which eventually decreases the distance. In Fig. 3, the positions of sensor node are being updated thereby at initial state, the moving distance of s_1 and s_2 is $d_1 + d_2$ and after the displacement, it will be updated to $d_3 + d_4$ as depicted in Fig. 4. It is worth mentioning that $d_1 + d_2 > d_3 + d_4$, therefore achieving the intended positions, the moving distance of s_1 and s_2 can be confined but no change will be occurred in coverage area but the area coverage distance rate will be extended. The sensor nodes that eager to update their moving position will be substitute with the moving position of the nodes which are stationary and does not require to move further. This step can prevent the nodes to make unnecessary and longer movement. In case the node does not possess sufficient energy while reaching at intended position, the other surrounding node will surrogate the liability. Initially, the node s_1 and s_2 tries to shift their positions with Intended-SP1 and thereupon establishes the desired link.

In fact, the distance between node s_1 and SP1 is greater than that one of its neighouring node. Similarly, node s_2 has a longer path to access the intended-SP1 position and meanwhile there appears another node in its surrounding which is much closer.

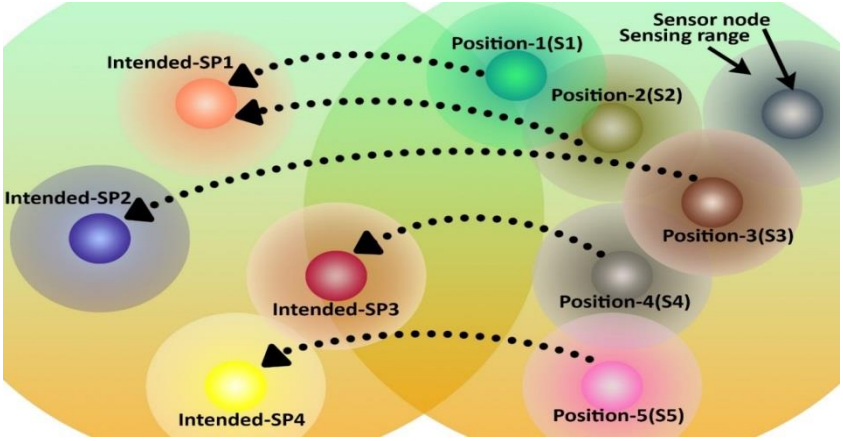
Both sensor nodes s_1 and s_2 will be hiatus to shift their positions depicted by the Fig. 3, but will change their positions according to the Fig. 4, i.e., s_1 will avail the new intended-sp1 position with distance d_3, and s_2 will be shifted to surrounding node position (intended-sp3) under distance d_4.

All this happens because $d_3 < d_1$ and $d_4 < d_2$ therefore the proposed DECM algorithm shrewdly decides which sensor node should be moved to what positions in accordance to the distance metric. This change will not affect the coverage range of the network and does not impel the rest of the nodes to move in the queue. Eventually, an average moving distance of the node node are reduced which enhance the coverage area distance range.

C. Affective parameters

The DECM, explains the impact of some effective elements such as, loudness, pulse emission rate and maximum frequency and wavelength which directly influence the performance and diversity the sensor coverage range.

• Loudness: It is a relative quality of sound or the characteristics that characterizes noises varying from silent to strong [28]. The sensing coverage range of the nodes are directly affected by the ambient forms of varying loudest. The proposed DECM algorithm handles the loudness by adjusting radius of the sensing range in the coverage area. It is also called a perceptual effect of the intensity of the sound. There are several models for this parameter estimation that aim to process numerically a loudness degree approximation dependent on the sound's objective characteristics.



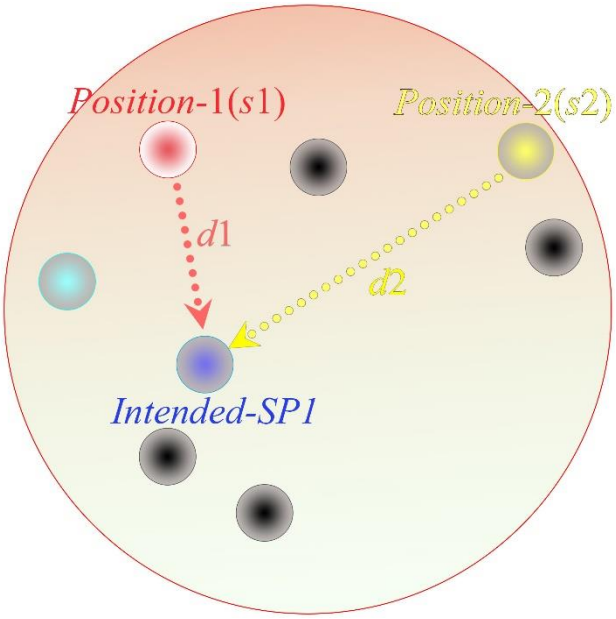


Fig. 3: Locating intended position mechansim.

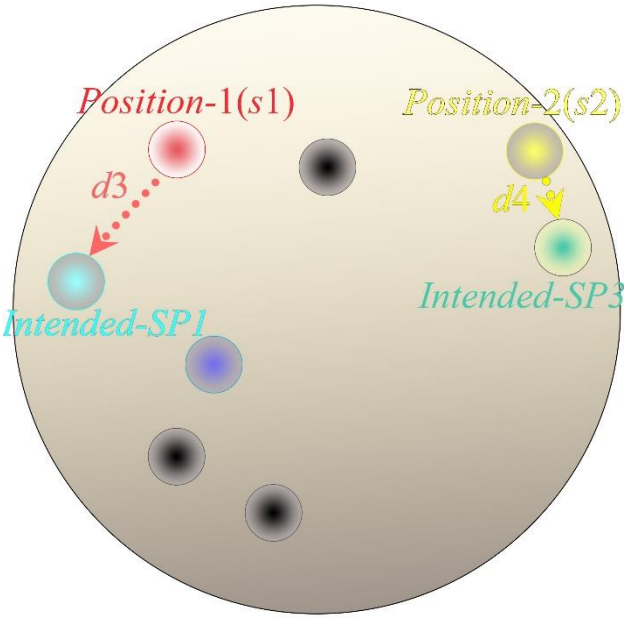


Fig. 4: Sensor nodes reaching at target position.

Algorithm 1. DECM position displacement mechanism

1: **Input distance**($d1$, $d2$, $d3$, $d4$) // distance metrics

2: **Procedure InitialPositions**($S1$, $S2$, $S3$, $S4$, $S5$) // sensor nodes deployed at initial positions

3: **Procedure IntendedPositions**($SP1$, $SP2$, $SP3$, $SP4$) // sensor nodes accessing new positions

4: **for** F(s) = {$S1$, $S2$, $S3$.....S_n}

5: ∀ $F(s)$, **if** $d3 > d1$ *AND* $d4 > d2$ // distance metric computation

6: **Compute** distance using Eq. (14)

7: S1 ← s$p1$, S2 ← s$p3$, // changing positions according to the displacement criteria

8: **if** (Rarea(S) reduces) $S1$ = Position; moved[S] = false;

9: **if** ($S1 \neq SP1$) and (moved[$S1$]) and (moved[$S2$]) and ($d1 + d2 > d3 + d4$)

10: $S1 \leftrightarrow SP1$, $S2 \leftrightarrow SP3$

11: **endfor**

12: **endif**

13: **end Procedure**

• Pulse emission rate: A pulse is a bit of disturbance emitted by any transmission source and travels through a medium [29]. The free end and fixed end pulses sometimes collide with other pulses which may narrow down the performance of the system. The DECM control the pulse emission rate by predicting the distance between sensing range of the nodes. The pulse emission rate for each sensor nodes is calculated as the total number of pulses detected divided by total duration of the number of individuals in the coverage area.

• Frequency and wavelength: In order to locate the intended target position, the shrewd adjustment in frequency and wavelength is often made by the DECM. Sometimes, for a shorter distance the frequency level might be shrinker from previous level.

• Group size(n): The number of sensor node appears in a couplet are referred as group size [30]. These nodes leys at varying positions and communicate simultaneously with surrounding nodes to access the

intended position. In fact, DECM makes shrewd decision in selection of the intended position from initial position by considering the distance metric between the sensor nodes at new position.

• Maximum iteration: The number of times the transmission rounds take place to achieve the desired position is known as iterations [31]. For a denser network, this might possible within a few iterations the target position can be achieved [32], but in case a sparse condition [33], the DECM decides about the number of iterations to be needed to complete the task.

• Step length estimation: A point indicates the initial position of sensor node and the number of total points going to record till reaching at the intended position is known as length of the steps [34]. The most common example of step length from real experience is a pedestrian positioning.

Simulation Results

In order to validate the performance of sensor nodes based on DECM algorithm, the simulation trials are conducted using MATALAB R2016a [12]. The performance among DECM, FOA, PSO and ACO are carried out using simulation setup parameters given in Table 2, in term of coverage range, computation time, standard deviation and network energy diminution. Continuing, nearbout 65 sensor nodes were deployed randomly in the monitoring area of size 60 × 60 m2. The initial and final sensor node deployment is illustrated in Fig. 6 and 7. As the transmission begins, it can be clearly understanding that node deployment based on (DECM) has minimum redundancy and is utmost uniform as compared to node deployment by the FOA mechanism.

Table 2: Simulation setup

Contents	Setting value
Deployment area	60 x 60 m^2
Number of sensor nodes	65
Grid point	0.4 m*0.4 m
Group size	20
Sensing radius,	5 m
Maximum iterations	25
Loudness	0.5
Pulse emission rate	0.5
f_{min}	0
f_{max}	2

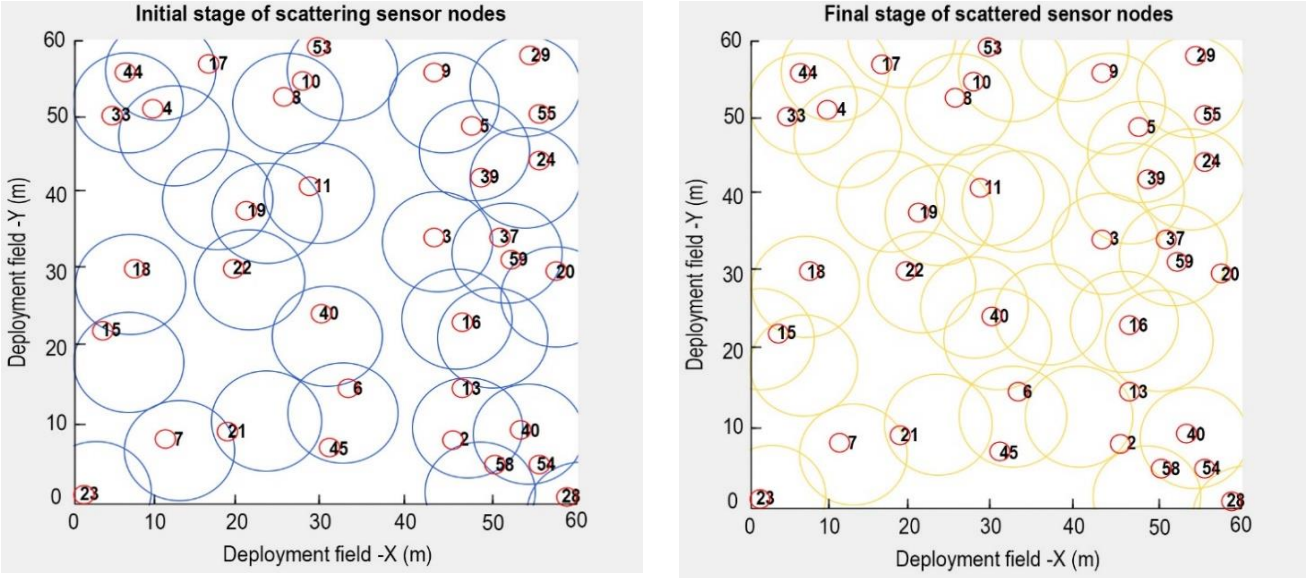
Table 3 signifies the influence of pulse emission rate (r) on coverage of sensor nodes. The value of r changes from 0.1 to 1 whereas value of other parameters such as loudness, maximum frequency and sensing radius is kept constant to 0.5, 2 and 5 respectively. To beat the effect of arbitrariness [35],

the node mechanism is simulated 50 times and greatest value of coverage is picked every time. The maximum value of coverage after performing DECM is attained 93.54% at pulse emission rate of 0.9. As node moves towards respective target, they emit a greater number of pulses, therefore, the pulse emission rate will be high when sensor nodes move close to the range points [36].

Thereupon, value of pulse emission rate is kept to 0.9. Further to analyze the effect of loudness of the mechanism on the coverage rate of sensor nodes, the value of loudness (Ao) is varied from 0.1 to 1 while pulse emission rate (r) is set to 0.9 and value of other parameters such as is 0.5, sensing radius (rs) is fixed to 5 meters.

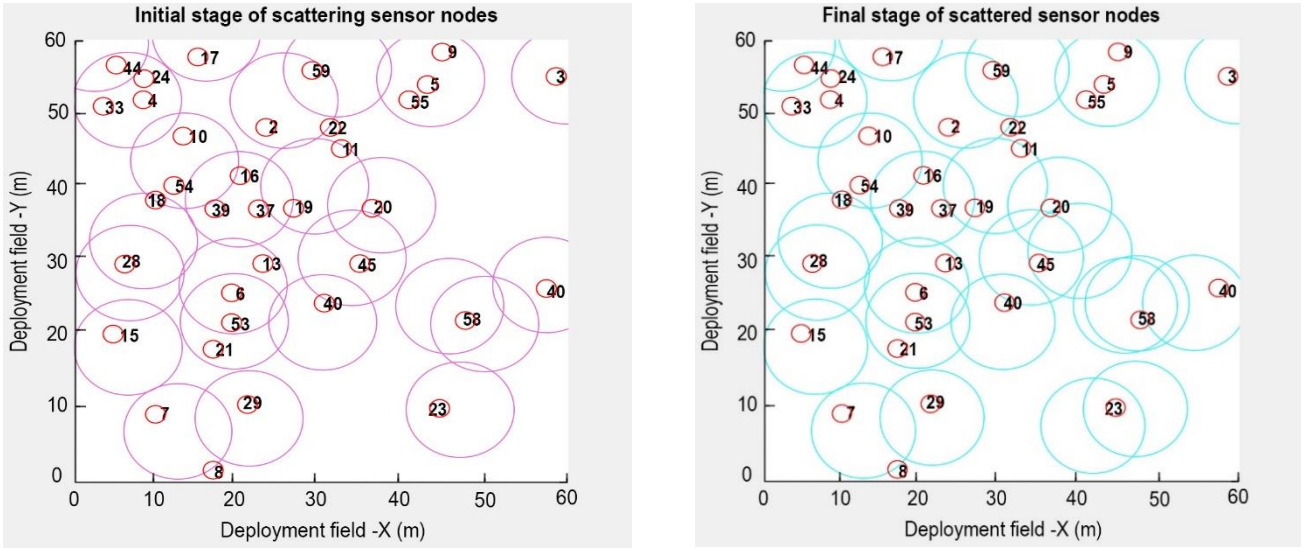
Table 4, shows the variations of loudness, initial and final coverage rate of nodes after implementing DECM. The DECM runs 50 times and best value of initial and final coverage range is selected. The coverage range after executing DECM has obtained highest 93.1% at 0.2 value of loudness. When sensor nodes getting near to the range point the intensity of emitted pulses is low, therefore loudness parameter should be kept low. Thereupon, the value of loudness parameter is fixed to 0.2. In addition to this Table 5, demonstrates the effect of maximum frequency (f_{max}) [37], on coverage; its value has been changed from 0.1 to 2. The constraints of the mechanism for instance the pulse emission rate, loudness and sensing radius are kept constant to 0.9, 0.2 and 5 respectively. For each variation of maximum frequency the proposed mechanism has been executed 50 times and supreme values of coverage before and after execution of the mechanism has been chosen.

The best value of coverage after implementing DECM is 93.31% when f_{max} is 1.3. Thus, the value of fmax is set to 1.3. To observe the impact of range points on coverage rate of nodes, value of range point has varied from 0.1 m*0.1 m to 1 m*1 m. The various simulation factors such as pulse emission rate, maximum frequency, sensing radius and loudness are kept constant to 0.9, 1.3, 5 and 0.2 respectively. In Table 6, for every value of coverage point DECM runs 50 time and uppermost values of coverage rate has been taken. The highest value of coverage rate of nodes is obtained after running DECM is 93.41% when range points are set to 0.6 m*0.6 m. Consequently; the range points have been kept constant to 0.6 m*0.6 m. Further, the sensing radius is varied from 1 m to 10 m.



(a) (b)

Fig. 6: (a) Initial and (b) final FOA node deployment mechanism.



(a) (b)

Fig. 7: (a) Start and (b) end of the node deployment process for DECM.

Fig. 8, signifies the variations of coverage range after applying DECM w.r.t. changes in the sensing radius of node. The parameters of DECM for example range points, loudness, pulse emission rate and maximum frequency are set as 0.6 m*0.6 m, 0.2, 0.9 and 1.3 respectively. It is clear from Fig. 8, as the sensing radius has increased, thereby coverage rate of sensor nodes is also increased and its value is 100% when the sensing radius is increased beyond 7 m, but there is trade-off between the sensing radius and cost, while sensing radius of node is increased the cost of sensor nodes also increased. The value of various constraints of DECM such as loudness, maximum frequency, sensing radius, pulse emission rate and range points are 0.2, 1.3, 6, 0.9 and 0.6 m*0.6 m respectively. To validate the performance of node deployment based PSO after setting above constraints values, the initial and final node deployment after executing are shown in Fig. 9. Thereupon, it can obviously be seen that node deployment based on DECM has lowest redundancy than PSO and FOA. To further demonstrates the effectives of coverage range curve for DECM compare to FOA for various iterations shown in Fig. 10. The iterations are varied from 0 to 500. The convergence speed of DECM is exorbitant as compared to FOA. The PSO converged arounds are 150 iterations whereas FOA converges around reached 350 iterations due to exploitation characteristics of the sensor nodes.

Table 3: Influence of pulse emission rate on coverage range

Pulse Emission (Hz)	Initial Coverage Range (%)	Final Coverage Range (%)
0.1	0.8	0.8929
0.2	0.8124	0.905
0.3	0.787	0.9077
0.4	0.8281	0.9041
0.5	0.8097	0.908
0.6	0.8202	0.9025
0.7	0.8208	0.9218
0.8	0.8167	0.9108
0.9	0.8537	0.9354
1	0.8314	0.9153

Table 4: Effect of loudness on coverage range

Loudness (A_o) db	Initial Coverage Range (%)	Final Coverage Range (%)
0.1	0.8052	0.8931
0.2	0.8375	0.9291
0.3	0.8491	0.9056
0.4	0.8281	0.9107
0.5	0.8276	0.9167
0.6	0.828	0.9219
0.7	0.8273	0.9048
0.8	0.8308	0.9259
0.9	0.8343	0.9281
1	0.8169	0.9179

Table 5: Effect of f_{max} on coverage range

f_{max} (Hz)	Initial Coverage Range (%)	Final Coverage Range (%)
0.1	0.8492	0.8698
0.2	0.819	0.8433
0.3	0.8135	0.8359
0.4	0.8115	0.8327
0.5	0.831	0.8602
0.6	0.8186	0.8507
0.7	0.8196	0..8414
0.8	0.8211	0.8417
0.9	0.8499	0.8712
1	0.8369	0.8549
1.1	0.8298	0.8888
1.2	0.822	0.9053
1.3	0.8134	0.9331
1.4	0.7965	0.898
1.5	0.8116	0.91
1.6	0.8367	0.9279
1.7	0.8145	0.9169
1.8	0.8267	0.9132
1.9	0.8296	0.9147
2	0.8127	0.9078

Table 6: The impact of range on network coverage

Range points (m*m)	Initial Coverage Range (%)	Final Coverage Range (%)
0.1*0.1	0.8306	0.9203
0.2*0.2	0.7975	0.9006
0.3*0.3	0.8006	0.9106
0.4*0.4	0.8342	0.9132
0.5*0.5	0.8012	0.9056
0.6 *0.6	0.8451	0.9341
0.7*0.7	0.8052	0.9125
0.8*0.8	0.8135	0.9181
0.9*0.9	0.8142	0.9200
1*1	0.8240	0.9212

The DECM has achieved more coverage rate about 99.46% as compared to 93.37%, 88.33% of PSO and FOA. In order to overwhelm the effect of randomness DECM, mechanism optimization and fruit fly algorithm runs 15 times respectively.

The deployment results in terms of average coverage rate, standard deviation, best and worst coverage values for DECM, PSO and FOA are presented in Fig.11 to 13.

It can be seen that DECM has achieved the average coverage range about 98.29% as compared to 91.91%, 85.16% of PSO and fruit fly algorithm. Further the standard deviation of DECM is lower, therefore DECM is more stable as compared to FOA and PSO.

The best and worst coverage value for DECM are 99.46% and 97.31% as compared to 94.30% and 90.02%, 87.49% and 78.20% for PSO and FOA based on node deployment.

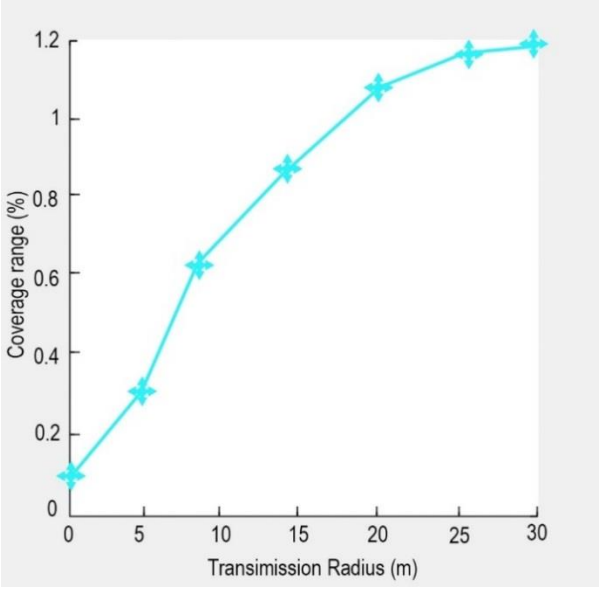
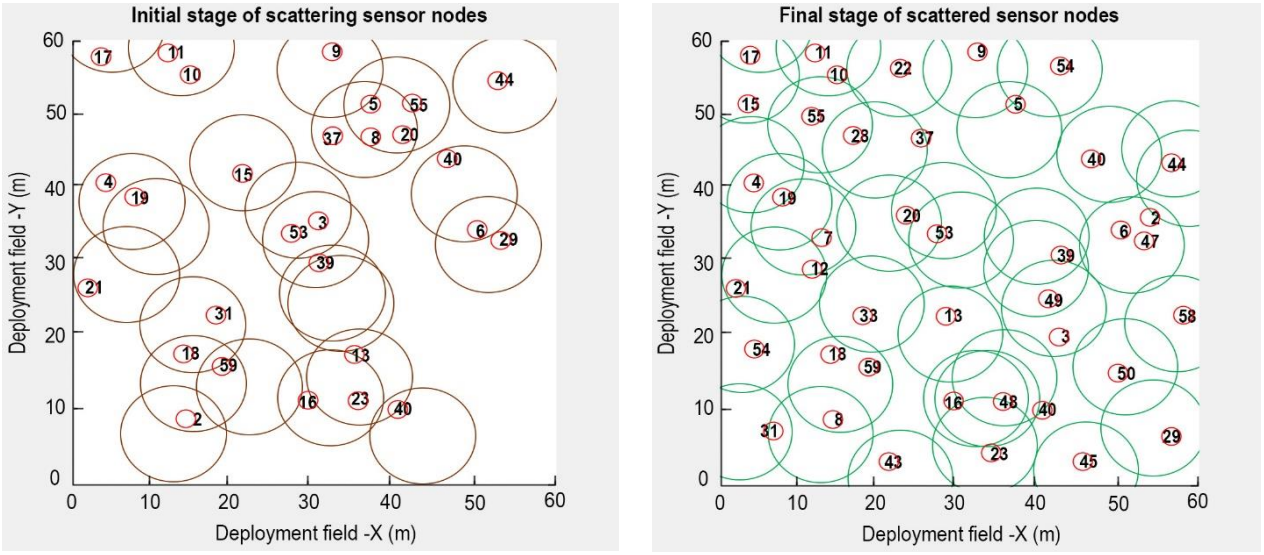
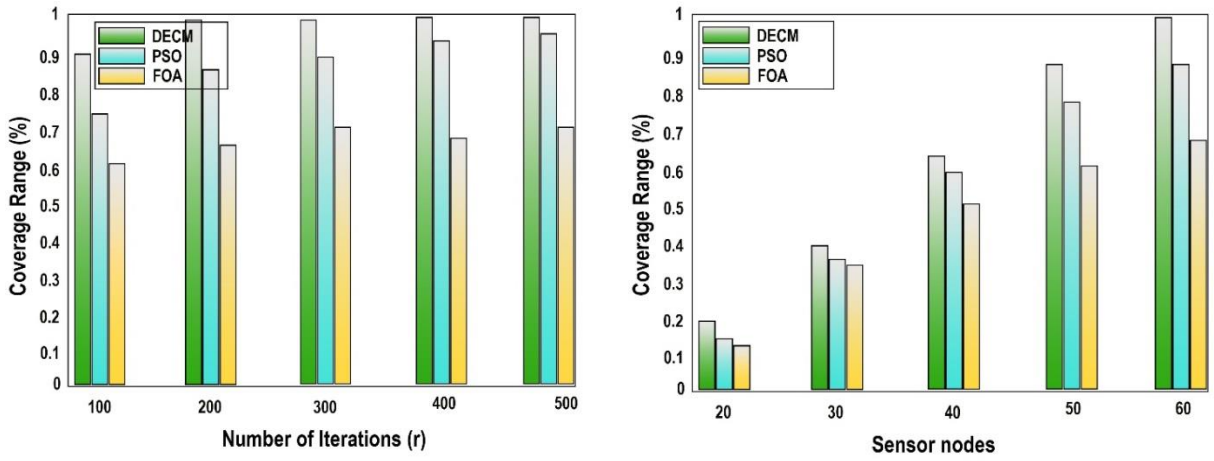


Fig. 8: DECM coverage range at various sensing radius .



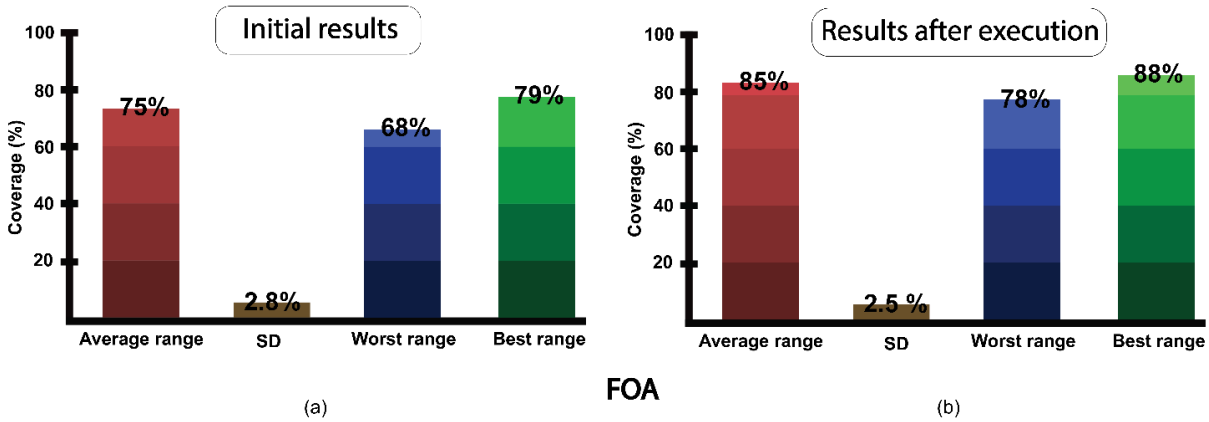
(a) (b)

Fig. 9: (a) Initial deployment of sensor nodes for PSO (b) Final deployment of sensor nodes for PSO.



(a) for various iterations (b) for varying number of sensor nodes

Fig. 10: Coverage range comparative analysis for DECM, FOA and PSO.



(a) (b)

Fig. 11: The coverage range statistics achieved by FOA (a) before and (b) after the execution process with significant changes in standard daviations.

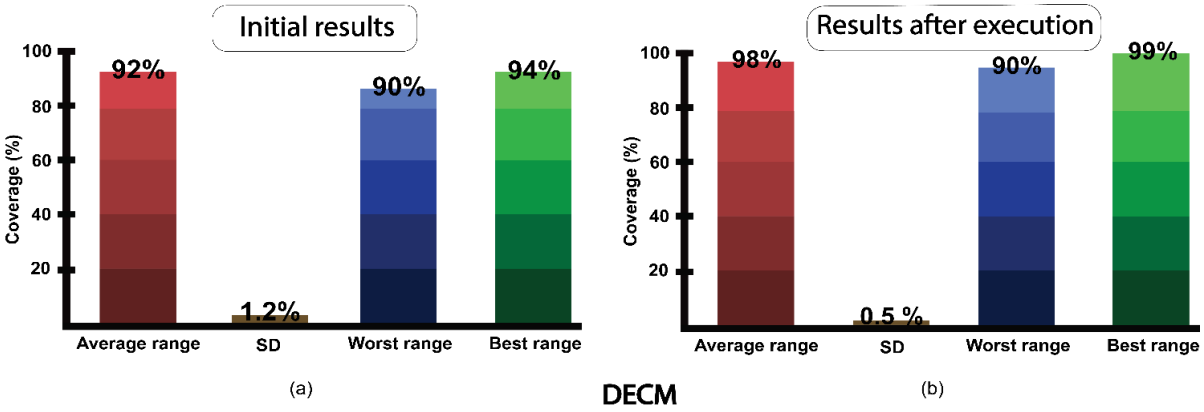


Fig. 12: The coverage range statistics achieved by PSO (a) before and (b) after the execution process with improved standard daviations.

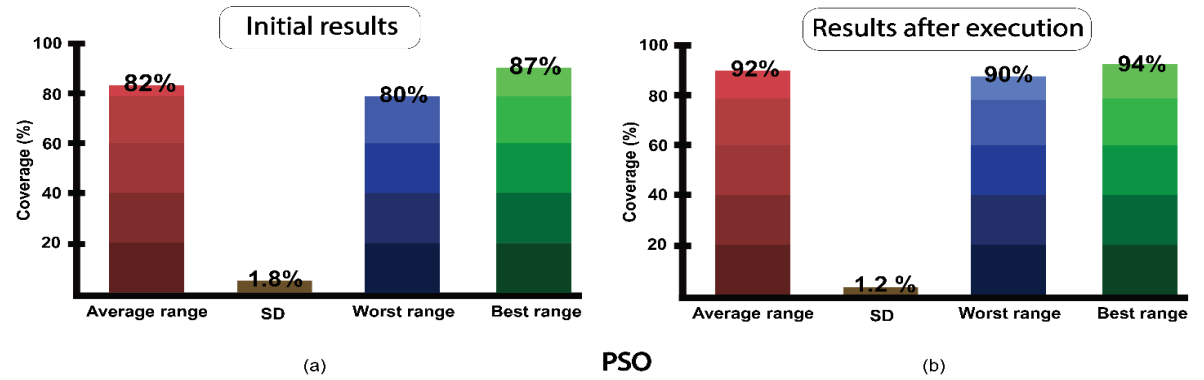


Fig. 13: The coverage range statistics achieved by DECM (a) before and (b) after the execution process with trivial standard daviations.

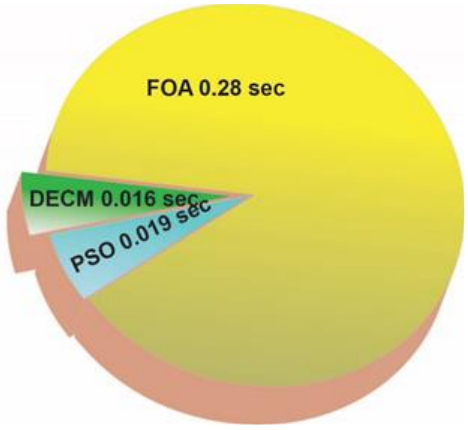


Fig. 14: The computation time statistics of FOA, PSO and DECM.

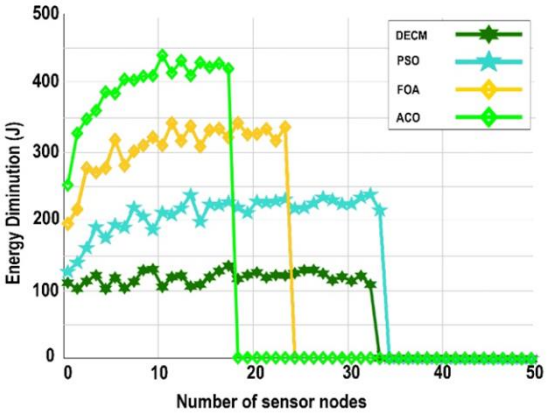


Fig. 15: Overall energy consumption by all algorithms.

Further the comparison of DECM, PSO and FOA in terms of computation time is ilustrated in Fig. 14. The computation time for DECM is lesser i.e. 0.016 seconds as compared to 0.019 seconds, 0.28 seconds for PSO and FOA. The DECM and PSO converges at 25 iterations whereas FOA converged at 500 iterations, therefore the speed of DECM and PSO is more and converges faster at earlier stage because of its exploitation feature as compared to fruit fly algorithm. During each transmission round, the overall energy diminution analysis is illustrated in Fig. 15. It can be seen

that between 20 and 32 nodes all algorithms going to die. The proposed DECM has consumed only 100 to 150 jule of energy approaching to relative position as compare to PSO, FOA and Ant Colony Optimization ACO. The consumed energy increases when

the coverage degrees required increase, since the sensor nodes require more energy to cover target positions and therefore it takes more energy for sensing and communication tasks.

Conclusion

Wireless sensor networks are severely facing the coverage issues therefore a shrewd coverage mechanism is presented in this study. The proposed algorithm Node Redeployment Shrewd Mechanism (DECM) has been designed to overcome the tribulations occurred due to the uncouth deployment of the sensor node which ultimately has great impact over network coverage range. The DECM functions in two phases, in first phase it searches the new intended node positions through Dissimilitude Enhancement Scheme (DES) and moves the node to new position. For second phase, the distance measurement between moving sensor node and the intended position is reduced and number of sensor movements are also being controlled sagaciously. This process is called Depuration. The analysis of various factors of DECM such as loudness, range points, emission rate and radius of nodes, and frequency have also been identified. The performance metrics of DECM has been obtained by conducting simulation test through MATLAB and meticulously compared with previous well-known algorithms FOA, PSO and ACO. The simulation results vouched that DECM has attained mean coverage range about 98.29% which is higher as compared to rest of the algorithms. The proposed DECM algorithm has appeared with higher coverage range and less computation time compared to all.

In future the various evolutionary optimization algorithms can be applied to solve the node deployment issues to enhance the coverage range phenomenon.

Author Contributions

Every author has equally contributed to accomplish the targeted results.

Acknowledgment

This work is completely self-supporting, thereby no any financial agency's role is available.

Conflict of Interest

All authors declare that there is no conflict of interests regarding the publication of this manuscript.

References

[1] M. Abazeed, N. Faisal, S. Zubair, A. Ali, "Routing Protocols for Wireless Multimedia Sensor Network: A Survey," 2013: 1-12, 2013.

[2] S. Ashraf, T. Ahmed, "Machine Learning Shrewd Approach For An Imbalanced Dataset Conversion Samples," J. Eng. Technol. JET, 11(1): 1-9, 2020.

[3] M. S. Aliyu, A. H. Abdullah, H. Chizari, T. Sabbah, A. Altameem, "Coverage Enhancement Algorithms for Distributed Mobile Sensors Deployment in Wireless Sensor Networks," Int. J. Distrib. Sens. Netw., 12(3): 9169236, 2016.

[4] S. Ashraf, Z. A. Arfeen, M. A. Khan, T. Ahmed, "SLM-OJ: Surrogate Learning Mechanism during Outbreak Juncture," Int. J. Mod. Trends Sci. Technol., 6(5): 162–167, 2020.

[5] F. Ait Aoudia, M. Gautier, M. Magno, O. Berder, L. Benini, "A Generic Framework for Modeling MAC Protocols in Wireless Sensor Networks," IEEEACM Trans. Netw., 25(3): 1489–1500, 2017.

[6] S. Ashraf, A. Ahmad, A. Yahya, T. Ahmed, "Underwater routing protocols: Analysis of link selection challenges," AIMS Electron. Electr. Eng., 4(3): 234–248, 2020.

[7] S. Ashraf, D. Muhammad, Z. Aslam, "Analyzing challenging aspects of IPv6 over IPv4," J. Ilm. Tek. Elektro Komput. Dan Inform., 6(1): 54–67, 2020.

[8] S. Ashraf, Z. Aslam, S. Saleem, S. Afnan, M. Aamer, "Multi-biometric Sustainable Approach for Human Appellative," CRPASE Trans. Electr. Electron. Comput. Eng., 6(3): 146–152, 2020.

[9] S. Goyal, M. S. Patterh, "Flower pollination algorithm based localization of wireless sensor network," in Proc. 2015 2nd International Conference on Recent Advances in Engineering Computational Sciences (RAECS): 1–5, 2015.

[10] S. Ashraf, Z. Aslam, A. Yahya, A. Tahir, "Underwater Routing Protocols Analysis of Intrepid Link Selection Mechanism, Challenges and Strategies," Int. J. Sci. Res. Comput. Sci. Eng., 8(2): 1–9, 2020.

[11] S. Ashraf, M. Gao, Z. Mingchen, T. Ahmed, A. Raza, H. Naeem, "USPF: Underwater Shrewd Packet Flooding Mechanism through Surrogate Holding Time," Wirel. Commun. Mob. Comput., 2020: 1–12, 2020.

[12] S. Ashraf, A. Raza, Z. Aslam, H. Naeem, T. Ahmed, "Underwater Resurrection Routing Synergy using Astucious Energy Pods," J. Robot. Control JRC, 1(5): 173-184, 2020.

[13] S. Ashraf, S. Saleem, A. H. Chohan, Z. Aslam, A. Raza, "Challenging strategic trends in green supply chain management," Int. J. Res. Eng. Appl. Sci. JREAS, 5(2): 71–74, 2020.

[14] S. Ashraf, T. Ahmed, S. Saleem, Z. Aslam, "Diverging Mysterious in Green Supply Chain Management," Orient. J. Comput. Sci. Technol., 13(1): 22–28, 2020.

[15] S. Ashraf, T. Ahmed, A. Raza, H. Naeem, "Design of Shrewd Underwater Routing Synergy Using Porous Energy Shells," Smart Cities, 3(1): 74–92, 2020.

[16] S. Ashraf, T. Ahmed, "Dual-nature biometric recognition epitome," Trends Comput. Sci. Inf. Technol., 5(1): 008–014, 2020.

[17] Q. Zhang, M. Fok, "A Two-Phase Coverage-Enhancing Algorithm for Hybrid Wireless Sensor Networks," Sensors, 17(12): 117, 2017.

[18] J. Li, B. Zhang, L. Cui, S. Chai, "An Extended Virtual Force-Based Approach to Distributed Self-Deployment in Mobile Sensor Networks," Int. J. Distrib. Sens. Netw., 8(3): 417307, 2012.

[19] G. Wang, G. Cao, T. F. La Porta, "Movement-assisted sensor deployment," IEEE Trans. Mob. Comput., 5(6): 640–652, 2006.

[20] H. Mahboubi, A. G. Aghdam, "Distributed Deployment Algorithms for Coverage Improvement in a Network of Wireless Mobile Sensors: Relocation by Virtual Force," IEEE Trans. Control Netw. Syst., 4(4): 736–748, 2017.

[21] W.-T. Pan, "A new Fruit Fly Optimization Algorithm: Taking the financial distress model as an example," Knowl.-Based Syst., 26:

69–74, 2012.

[22] M. Marinaki, Y. Marinakis, "A Glowworm Swarm Optimization algorithm for the Vehicle Routing Problem with Stochastic Demands," Expert Syst. Appl., 46: 145–163, 2016.

[23] Y. Yoon, Y.-H. Kim, "An Efficient Genetic Algorithm for Maximum Coverage Deployment in Wireless Sensor Networks," IEEE Trans. Cybern., 43(5): 1473–1483, 2013.

[24] L. Sun, X. Song, T. Chen, "An Improved Convergence Particle Swarm Optimization Algorithm with Random Sampling of Control Parameters," Journal of Control Science and Engineering, 2019: 1-1-11, 2019.

[25] "How to Calculate Euclidean Distance.", 2020.

[26] S. Ashraf, S. Saleem, T. Ahmed, Z. Aslam, D. Muhammad, "Conversion of adverse data corpus to shrewd output using sampling metrics," Vis. Comput. Ind. Biomed. Art, 3(19): 1-13, 2020.

[27] R. Storn, K. Price, "Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces," J. Glob. Optim., 11(4): 341–359, 1997.

[28] S. Balsamo, A. Marin, E. Vicario, Eds., New Frontiers in Quantitative Methods in Informatics: 7th Workshop, InfQ 2017, Venice, Italy, December 4, 2017, Revised Selected Papers, 1st ed. 2018 edition. New York, NY: Springer, 2018.

[29] S. Das, A. Biswas, S. Dasgupta, A. Abraham, "Bacterial Foraging Optimization Algorithm: Theoretical Foundations, Analysis, and Applications," in Foundations of Computational Intelligence Volume 3: Global Optimization, A. Abraham, A.-E. Hassanien, P. Siarry, and A. Engelbrecht, Eds. Berlin, Heidelberg: Springer: 23–55, 2009.

[30] M. Li, X. Du, X. Liu, C. Li, "Shortest Path Routing Protocol Based on the Vertical Angle for Underwater Acoustic Networks," Journal of Sensors, 2019: 1-14, 2019.

[31] R. Yongmao, L. Jun, S. Shanshan, L. Lingling, W. Guodong, Z. Beichuan, "Congestion control in named data networking – A survey," Comput. Commun., 86: 1–11, 2016.

[32] S. Ashraf, D. Muhammad, M. Shuaeeb, Z. Aslam, "Development of Shrewd Cosmetology Model Through Fuzzy Logic," J. Res. Eng. Appl. Sci., 5(3): 93–99, 2020.

[33] S. Ashraf, S. Saleem, T. Ahmed, "Sagacious Communication Link Selection Mechanism for Underwater Wireless Sensors Network," Int. J. Wirel. Microw. Technol., 10(4): 22–33, 2020.

[34] S. Ashraf, M. Gao, Z. Chen, S. Kamran, Z. Raza, "Efficient Node Monitoring Mechanism in WSN using Contikimac Protocol," Int. J. Adv. Comput. Sci. Appl., 8(11): 429-437, 2017.

[35] S. Ashraf, M. Gao, Z. Chen, H. Naeem, A. Ahmad, T. Ahmed, "Underwater Pragmatic Routing Approach Through Packet Reverberation Mechanism," IEEE Access, 8: 163091–163114, 2020.

[36] M. Safkhani, "Cryptanalysis of R2AP an Ultralightweight Authentication Protocol for RFID," J. Electr. Comput. Eng. Innov., 6(1): 111-118, 2018.

[37] S. Shams Shamsabad Farahani, "Congestion Control Approaches Applied to Wireless Sensor Networks: A Survey," J. Electr. Comput. Eng. Innov., 6(2): 129-149, 2018.

Biographies

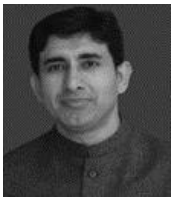
Shahzad Ashraf received B.E. degree in Computer Systems Engineering, and M.E. in Communication System and Networks from Mehran Engineering & Technology University, Jamshoro Pakistan in 2004, and 2014 respectively. Currently, he is a Ph.D. student at Hohai University Changzhou China. From 2005 to 2016, he served as an Assistant Professor at NFC Institute of Engineering and Technology Multan, Pakistan. His area of interest includes Wireless sensor communication, Underwater routing, Computer graphics and architecture, Computer Networks, Grid and distributed computing and Computer hardware. He is the active reviewer and member of technical committee of more than 50 renowned international journals and conference proceedings including IEEE and ACM.

Tauqeer Ahmed received B.SC degree in Computer Systems Engineering and M.S in Electronic Engineering from International Islamic University, Islamabad Pakistan in 2011, and 2015 respectively. He is currently pursuing the Ph.D. degree in Information and Communication Engineering with the College of Internet of Things of Engineering, Hohai University Changzhou Campus China from 2018 to 2022, he served as an IT officer at Punjab Information Technology Board Multan, Pakistan. His research interest includes Image processing, Signal processing, Underwater acoustic sensor network, Computer hardware and networks.

Zeeshan Aslam received B.E degree in Electrical (Computer System) Engineering from Bahauddin Zakariya University, Multan and M.S in Electrical (Power) Engineering from Institute of Southern Punjab Multan in 2015, and 2018 respectively. He is currently serving as a Site HSE Manager in Petroweld Oilfield Services Kurdistan region Iraq since 2019.He served as an Electrical and HSE Engineer in Volka Food International Multan Pakistan as an Electrical Engineer from 2015-2019.He also served as a Visiting Faculty in NFC Institute of Engineering and Technology Multan, Pakistan 2017-2018.

Durr Muhammad received B.Sc degree in Computer Systems Engineering from NFC-IET Multan, Punjab, Pakistan and M.E. in Electronic Systems Engineering from Mehran Engineering & Technology University, Jamshoro Pakistan in 2009 and 2014 respectively. He is currently working as Assistant Executive Engineer in Pakistan Steel Mills Karachi, Pakistan which is a country large Industrial Complex of Pakistan and previously worked as visiting Instructor in Dadabhoy Institute of Higher Education Karachi, Pakistan from Jan-2014 to 29th of October 2016.

Adnan Yahya received B.S. degree in Information Technology, from University of Sindh Jamshoro Pakistan in 2006 and MS degree in Computer and communication Network from Hamdard University in 2019 Karachi Pakistan, he served as an I.T Administrator at DUHS Dow University of Health Sciences, Karachi- Pakistan.



Muhammad Shuaeeb received BS degree in Computer and Information Technology from University of Sindh, Jamshoro in 2003, and MS in Computer and Communication Networks from Hamdard University, Karachi in 2016. He is serving as an IT Administrator at Dow University of Health Sciences, Karachi, Pakistan since 2010, where he is looking after Datacenter operations, Cloud services, Network security and designing, and Domain administration. His area of interest includes Cloud computing, Datacenter management, wireless communication, Network security, Identity and access management, Server hardware and networks, artificial intelligence, computer architecture, Data Storage devices, HA management and Virtualization.

How to cite this paper:
S. Ashraf, T. Ahmed, Z. Aslam, D. Muhammad, A. Yahya, M. Shuaeeb, "Depuration based Efficient Coverage Mechanism for Wireless Sensor Network," Journal of Electrical and Computer Engineering Innovations, 8(2): 145- 160, 2020.

DOI: 10.22061/JECEI.2020.6874.344

URL: http://jecei.sru.ac.ir/article_1453.html

Journal of Electrical and Computer Engineering Innovations
(JECEI)
Journal homepage: http://www.jecei.sru.ac.ir

Research paper

Quantitative Assessment of Transformation Based Satellite Image Pan-sharpening Algorithms

F. Tabib Mahmoudi *, ***A. Karami***

Department of Geomatics, Faculty of Civil Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran.

Article Info

Article History:
Received 03 September 2019
Reviewed 12 November 2019
Revised 17 January 2020
Accepted 21 April 2020

Keywords:
Pan-sharpening
Quantitative analysis
Spectral deviation
Transformation based method
Satellite imagery

*Corresponding Author's Email Address:
fmahmooudi@sru.ac.ir

Abstract

Background and Objectives: Pan-sharpening algorithms integrate the spectral capabilities of the multispectral imagery with the spatial details of the panchromatic one to obtain a product with confident spectral and spatial resolutions. Due to the large diversities in the utilized pan-sharpening algorithms, occurring spatial and spectral deviations in their results should be recognized by performing the quantitative assessment analysis.

Methods: In this research, the pan-sharpened images from PCA, IHS, and Gram-Schmidt transformation based algorithms are evaluated for the multi-spectral and panchromatic images fusion of Landsat-8 OLI sensor (medium scale resolution satellite) and WorldView-2 (high-resolution satellite). Quantitative analysis is performed on the pan-sharpened products based on the Per-Pixel Deviation (PPD) measure for spectral deviation analysis and high-pass filter and edge extraction measures for analyzing the spatial correlations. Moreover, entropy and standard deviation quantitative evaluation measures are also utilized based on the pan-sharpened image content.

Results: Quantitative analysis represents that increasing the spatial resolution of the utilized remote sensing data has direct impacts on the spectral, spatial, and content-based characteristics of the generated Pan-sharpened products. Gram-Schmidt transformation based pan-sharpening method has the least spectral deviations in both WorldView-2 and Landsat-8 satellite images. But, the amount of spectral, spatial and content-based quantitative measures of PCA and IHS are changing with various spatial resolutions.

Conclusion: it can be said that Gram-Schmidt pan-sharpening method has the best performance in both medium-scale and high-resolution data sets based on the spectral, spatial, and content quantitative evaluation results. The IHS pan-sharpening method has better performance than the PCA method in Landsat-8 OLI data. But, by increasing the spatial resolution of the data, PCA generates pan-sharpened products with better spectral, spatial, and content based quantitative evaluation results.

Introduction

Due to the increasing developments of the sensor technologies along with the modern information acquisition techniques, a large volume of remote sensing data with different spectral, spatial, and temporal characteristics has been provided to users. Various remotely sensed data has valuable information aspects that together can fully provide the information needed by researchers. Performing remotely sensed data fusion to generate new data that contains all the useful aspects

of information in each of the primary data has received much attention in image processing and pattern recognition [1]-[3]. One of the important applications of remote sensing data fusion is to increase the spatial resolution of multispectral imagery to the spatial details of panchromatic, which is called the pan-sharpening method [4], [5]. Pan-sharpening as a pixel level fusion has a wide variety of methods and algorithms. Therefore, researchers in various fields by categorizing the algorithms analyze the characteristics and advantages of them [2], [6]. Since, pan-sharpening algorithms are rapidly developing in some applications such as object recognition, image classification, and change detection, the efficiency of different integration algorithms can be considered as an essential need. The main necessity in pan-sharpening is that the integration algorithm should maintain as much information as possible in the input images. However, pan-sharpening algorithms usually cause some spatial and spectral distortions in the fused image. Therefore, a quantitative assessment is an essential process for content and distortion evaluation of the pan-sharpened image. Choosing the right pan-sharpening algorithm requires a quantitative evaluation of the results of different integration methods. The main objective of image fusion quantitative assessment is to obtain a quantitative estimate of the quality of the resulting image, which requires the definition of an appropriate metric. According to widely use and rapid developments of pan-sharpening algorithms in the satellite image processing and pattern recognition contributions, quantitative assessment metrics are still important in open research topics [3]-[5], [7]-[11]. Table 1 summarizes the investigated literature review in this research into two categories; pan-sharpening performance analysis, and analyzing the efficiencies of quantitative assessment metrics. In this study, the results of Intensity-Hue-Saturation (IHS), Principal Component Analysis (PCA), and Gram-Schmidt transformation based pan-sharpening methods are evaluated based on spectral and spatial reference-based quantitative assessment measures. Moreover, the results are compared with the quantitative assessment based on entropy and standard deviation measures that don't need reference images. The reason for choosing the above-mentioned transformation based pan-sharpening methods is that these methods are available in most common image processing software and in most cases, researchers perform these pan-sharpening methods to increase the spatial resolution of remote sensing images in various applications as a pre-processing step.

On the other hand, various types of remote sensing data with a high variety of spatial resolutions are utilized in different applications.

Table 1: Summarizing the literature review of this paper

Category	Research Topic	Publication Date [Ref]
Pan-sharpening performance analysis	PCA & Wavelet transformation based pan-sharpening assessment (for MRI images).	2010 [3]
	Comparing four pan-sharpening methods with a no-reference assessment (on IKONOS & WV2).	2014 [8]
	Spectral and spatial comparison of twelve pan-sharpening algorithms.	2020 [9]
	Quantitative analysis of a new pan-sharpening algorithm pulse coupled neural network.	2019 [10]
	Quantitative analysis of a new pan-sharpening algorithm based on IHS and wavelet.	2020 [11]
Quantitative assessment efficiency analysis	Introducing a new perceptual quality assessment of pan-sharpened images.	2019 [4]
	Performance comparison of pan-sharpening assessment measures.	2012 [5]
	Critical review of quality assessment protocols in pan-sharpening.	2019 [7]

Therefore, in this study, the results of PCA, IHS, and Gram-Schmidt pan-sharpening algorithms are applied to two types of remote sensing data with different spatial resolutions. Landsat-8 satellite imagery was used as a representative of the medium-scale spatial resolution data, and WordView-2 imagery was used as a representative of high spatial resolution data.

The main objective and novelty of this research with regard to the literature review is on performing assessment in dual aspects;

1) Comparing the efficiencies of well-known and common transformation based pan-sharpening algorithms in most of the image processing software.
2) Investigating the effects of increasing the spatial resolution of the commercial satellite images in pan-sharpening results.

The other point that can be considered as the novelty of this research is multi-modal conclusion based on the spatial, spectral and content quantitative assessment of the pan-sharpened products. In other words, this study in addition to compare the capabilities of PCA, IHS, and Gram-Schmidt transformation based pan-sharpening methods from spectral, spatial, and content points of

view; also analyze the effect of increasing the spatial resolution of remote sensing data on the generated pan-sharpened products.

Transformation Based Pan-sharpening Algorithms

The use of transformation based pan-sharpening methods is based on the fact that after transforming input images to a new space, by applying the appropriate rules, the images are merged and finally the result is obtained by applying the inverse transformation to the image space [12]-[13]. The important point in such image fusion methods is that to perform pan-sharpening, transformation is usually applied only to the multispectral image and then in the new space, the panchromatic image is replaced with one of its parameters. In this way, a multi-spectral image with the higher spatial resolution is generated, which is returned to the image space by applying the inverse transformation on it. According to the different natures of various transformation based pan-sharpening methods, Intensity-Hue-Saturation (IHS), Principal Component Analysis (PCA), and Gram-Schmidt are described in the following sections of this paper.

A. Principal Component Analysis

In the principal component analysis process, statistically correlated variables are converted to non-correlated variables and a compressed and optimal description of the input data is provided. Assuming an input image with dimensions of *M * N*, PCA method seeks to find a basic orthonormal function *W = (W1, W2,…, Wd) d << MN* so that the desired image can be displayed by a linear combination of these basic vectors [14], [15]. If the principal components of an image are in descending order in the PCA transformation space, the first component has the highest amount of variance in the image and is considered as the parameter containing the spatial detail of the image. This feature has led to the potential for the use of PCA transformation to integrate images with the aim of pan-sharpening [16], [17]. PCA pan-sharpening method has the following steps:

- **Step1:** Performing PCA transformation on the multispectral input image.
- **Step 2:** Histogram matching of the panchromatic image with the first principal component of the multispectral image.
- **Step 3:** Replacing the first component of the multispectral image in the PCA transformation space with the panchromatic image after histogram matching.
- **Step 4:** Applying the inverse PCA transformation to the new multispectral image to transform it into the original primary space.

B. HIS

Intensity, Hue, Saturation (IHS) is an image display system that includes the three main parameters intensity, shape, and saturation. Intensity as the first parameter of the IHS system relates to the overall brightness of a color. The parameter Hue corresponds to the color wavelength and the third parameter Saturation indicates the color purity. Since the display of color images in the IHS display system is based on the parameters of human color perception, the use of image transformation method from RGB to IHS color space in the preliminary stages of image processing is very common. The IHS color transformation method effectively separates the weight of the spatial information of the image (Intensity) from the weight of its spectral information (Hue, Saturation). Therefore, this transformation method is useful for performing pan-sharpening by replacing the panchromatic image with the Intensity parameter of the multispectral image in the IHS color space [18], [19].

C. Gram-Schmidt

Gram-Schmidt (GS) transformation is a common method for having orthogonal basic vectors of a space. A matrix or an image can also be used as input in GS conversion. Assuming that the set S = {v_1, v_2,…, v_n} are the vectors of the orthogonal base of the interior multiplication space V, each vector $w \in V$ can be shown as the linear combination of the base vectors S.

$$w = \frac{\langle w, v_1 \rangle}{\|v_1\|^2} v_1 + \frac{\langle w, v_2 \rangle}{\|v_2\|^2} v_2 + \ldots + \frac{\langle w, v_n \rangle}{\|v_n\|^2} v_n \quad (1)$$

Now, if we assume that {u_1, u_2,…, u_n} is the desired base in the interior multiplication space of V, then using the Gram-Schmidt algorithm we can form the orthogonal base {v_1, v_2,…, v_n}:

$$\mathrm{Proj}_v(u) = \frac{\langle u, v \rangle}{\|v\|^2} v \quad (2)$$

The Gram-Schmidt transformation method has been used successfully to integrate images with the aim of pan-sharpening. In this method, unlike IHS color space transformation, there is no limitation on the number of spectral bands that can be processed in pan-sharpening. Gram-Schmidt pan-sharpening method has the following steps:

- **Step1:** Restore a panchromatic image with the low spatial resolution
- **Step 2:** Applying the GS conversion on the low spatial resolution panchromatic image and multispectral imagery. Here, the low spatial resolution panchromatic image is the first band of GS.
- **Step 3:** Replacing the main panchromatic image (high spatial resolution) with the first band of GS transformation. For this purpose, it is necessary to first match the mean and standard deviation of the high

spatial resolution panchromatic image with the first band of GS conversion.

Pan-sharpening quantitative assessment analysis

As pan-sharpening quantitative assessment methods are based on suitable metrics, and according to the nature of the designed metrics, these methods can be categorized into one of the following two groups [20]:

- Reference-based quantitative assessment methods. In these methods, the information contained in the input images to the fusion algorithm is used as a reference for quality evaluation of the pan-sharpened image. As the main objective of performing pan-sharping is to increase the spatial resolution of a multispectral image by combining it with the spatial detail of a panchromatic image, for quantitative assessment of pan-sharpened image, it can be spectrally compared with the input multi-spectral image. Moreover, the spatial characteristic of the pan-sharpened image is also evaluated by the input panchromatic image.
- Quantitative assessment methods those metrics are worked without the need for a reference image and only based on the information contained in the pan-sharpened image.

In this paper, the results of applying the PCA, IHS, and Gram-Schmidt transformation based pan-sharpening methods on the panchromatic and multi-spectral WorldView-2 and Landsat-8 images are evaluated quantitatively. For this purpose, quantitative evaluation methods based on spectral and spatial comparison of the pan-sharpened image with input images to the algorithm are utilized. Also, no-reference based quantitative assessment methods are used for evaluating only based on the information content of the pan-sharpened image. In the following sections, after introducing the data sets used in this research, first, each of the utilized quantitative evaluation metrics will be introduced and then, the quantitative assessment results of the pan-sharpened products will be presented and discussed.

Data Sets

For quantitative assessment of the transformation based pan-sharpening methods, panchromatic and multispectral images of WorldView-2 and Landsat-8 OLI sensor are utilized. The WorldView-2 satellite has a panchromatic image with half a meter spatial resolution. Moreover, the multispectral sensor of the WorlView-2 has 8 spectral bands with 2.4 meters spatial resolution. Thus, the generated pan-sharpened image from WorldView-2 panchromatic and multispectral sensors has 8 spectral bands with half a meter spatial resolution.

Landsat-8 OLI sensor has a panchromatic image with 15 meters spatial resolution and a multispectral image with 7 spectral bands and 30 meters spatial resolution. The generated pan-sharpened image from the Landsat-8 OLI sensor has 7 spectral bands with 15 meters spatial resolution. Detail characteristics of the WorldView-2 and Landsat-8 OLI data are depicted in Table 2.

Table 2: The characteristics of WorlView-2 and Landsat-8 OLI dat

Landsat-8 OLI			WorldView-2	
No	Band	Wavelength (μm)	Band	Wavelength (μm)
1	Blue	0.45-0.52	Coastal	0.40-0.45
2	Green	0.52-0.60	Blue	0.45-0.51
3	Red	0.63-0.69	Green	0.51-0.58
4	NIR	0.77-0.90	Yellow	0.59-0.63
5	SWIR1	1.55-1.75	Red	0.63-0.69
6	TIR	10.4-12.5	Red Edge	0.70-0.74
7	SWIR2	2.09-2.35	NIR1	0.77-0.895
8	PAN	0.52-1.90	NIR2	0.86-0.95

The main reason for performing a quantitative assessment on the pan-sharpened images of these two types of satellite data is their differences in spatial resolution. The WorldView-2 pan-sharpened image has a high spatial resolution and Landsat-8 OLI pan-sharpened image has a medium-scale spatial resolution. Both of the utilized data are taken from the same urban area in San Francisco. The pan-sharpened WorldView-2 satellite image has 9270*10140 pixels with 0.5*0.5 square meters and the pan-sharpened Landsat-8 satellite image has 348*394 pixels with 15*15 square meters.

Spectral Metrics

As it is mentioned in previous sections, for spectral quantitative assessment of the pan-sharpened image, spectral metrics are described those use the input multispectral image as a reference. Per Pixel Deviation (PPD) is the utilized quantitative measure in this research for evaluating the spectral distortions in the pan-sharpened image.

PPD metric: This metric is used to calculate the amount of spectral deviations between each pixel of the pan-sharpened image and the initial multispectral image to the fusion algorithm. Quantitative assessment based on the PPD metric includes the following steps (see Fig.1):

- **Step 1:** Decreasing the spatial resolution of the pan-sharpened image as equal to the spatial resolution of the input multispectral image.
- **Step 2:** Pixel by pixel differencing between the original multispectral image and the pan-sharpened image with reduced spatial resolution.
- **Step 3:** Determining the average differences for each pixel based on the gray values.

The results of performing the PPD metric as spectral quantitative assessments on the pan-sharpened images of Landsat-8 OLI and WorldView-2 are depicted in Table 3. According to the results, the PPD metric shows the

best spectral matches (the least spectral deviations) between the multispectral image and the pan-sharpened one that is generated by the Gram-Schmidt method. For PCA and IHS pan-sharpening methods, there are more spectral deviations in the generated pan-sharpened images. Since the first principal component in PCA transformation, in addition to spatial details, contains some spectral information, totally replacing it with a panchromatic image causes the loss of some spectral information in the resulting pan-sharpened product.

Table 3: Spectral quantitative assessment of pan-sharpened images based on PPD

Methods	Landsat-8 OLI	WorldView-2
PCA	28.2921	0.0663
IHS	15.1158	0.0766
Gram-Schmidt	**0.8971**	**0.0313**

This leads to spectral distortions and variations in the pan-sharpening results from the PCA transformation based fusion method.

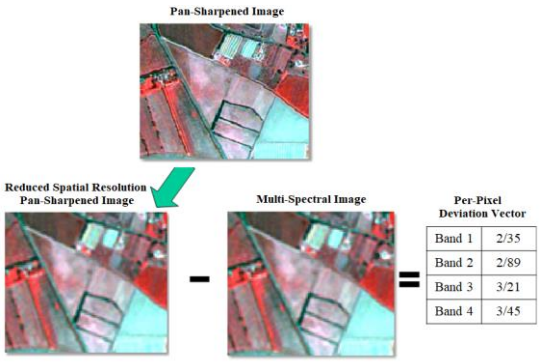


Fig. 1: Per-Pixel Deviation (PPD) quantitative assessment process.

IHS method has the limitation of the number of spectral bands to process. This method applies only to the three spectral bands of the generated pan-sharpened images. Therefore, more spectral deviations may occur.

Spatial Metrics

To spatial quantitative assessments of the pan-sharpened image, some metrics are described those use the input panchromatic image to the fusion algorithm, as reference. The high-pass filter and edge extraction metrics are utilized in this research for the spatially quantitative assessment of the pan-sharpened images from WorldView-2 and Landsat-8 OLI satellite sensors.

High-pass filter metric: As we know, by applying a high-pass filter to the image, its high frequencies can be extracted. Thus, in pan-sharpening algorithms, it is also possible to apply high-pass filters to the panchromatic and the resulting pan-sharpened images for spatially quantitative evaluation of the pan-sharpened algorithm. In this metric, first, a suitable high-pass filter is applied to the input panchromatic image and the generated pan-sharpened image to obtain the high frequencies of these two images. Then, the correlation coefficient between the filtered panchromatic image and each of the filtered bands of the pan-sharpened image is calculated. The greater correlation between the high frequencies of the pan-sharpened image and the high frequencies of the input panchromatic image means that more spatial details of the panchromatic image have been transferred to the pan-sharpened one.

Edge extraction metric: Another spatially quantitative assessment method of pan-sharpening algorithms is those using edge extraction operators such as Canny for edge extraction from the pan-sharpened image and the panchromatic image entering to the fusion algorithm (Fig. 2).

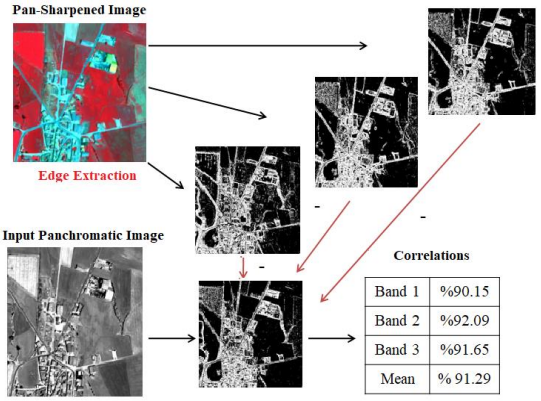


Fig. 2: Edge extraction quantitative assessment process.

The more similar the edge information extracted from the pan-sharpened image is to the edges of the input panchromatic image, the more successful the pan-sharpening algorithm has been in transmitting the spatial detail of the input panchromatic image to the pan-sharpened one.

As it is clear in Table 4, both of the edge extraction and high-pass filter metrics indicate the most spatial correlation between Gram-Schmidt pan-sharpened results and the input panchromatic image of Landsat-8 and WorldView-2.

No-reference Metrics

In the group of quantitative assessment methods for evaluating the quality of pan-sharpening algorithms, techniques have also been presented that do not require the use of a reference image in their proposed metrics. In this group of techniques, only using the information extracted from the pan-sharpened image, the quantitative evaluation of the pan-sharpening algorithm can be performed. In the following sub-sections, standard deviation and entropy are described as the well-known no-reference quantitative analysis measures.

Standard Deviation: This criterion performs the quantitative evaluation of the pan-sharpened image without needing reference images and only by obtaining the difference between the pan-sharpened image pixels'

gray values and its average value. According to this evaluation method, the larger the standard deviation, the greater the deviations of the pan-sharpened pixels' gray values from the average, and as a result, the distortion created in the result of the pan-sharpening algorithm is greater. As another description of this criterion, it can be stated that the standard deviation reflects the contrast of the information in the pan-sharpened image.

$$SD = \sqrt{\frac{1}{n-1}\sum_{i=1}^{M}\sum_{j=1}^{N}(I_f(i,j)-\mu)^2} \quad (3)$$

In (3), n is the number of spectral bands and μ is the pan-sharpened image average value.

Table 4: Spatial quantitative assessment of pan-sharpened images

Methods	Landsat-8 OLI		WorldView-2	
	Edge Extraction (%)	High-Pass Filter (%)	Edge Extraction (%)	High-Pass Filter (%)
PCA	74.02	29.386	98.18	92.82
IHS	84.89	73.02	89.58	92.89
Gram-Schmidt	**86.86**	**96.72**	**99.34**	**92.90**

Entropy: The entropy criterion is used to evaluate the information contained in the pan-sharpened image. The greater the information contents in the image, the greater the numerical entropy value. Entropy is sensitive to the noise of the image. In (4), h_{If} is the probabilities of the pixels' gray values of the pan-sharpened image I_f.

$$Entropy = -\sum_{i=0}^{L} h_{I_f}(i)\ .Log\ h_{I_f}(i) \quad (4)$$

The results of entropy and standard deviation quantitative evaluation metrics on the pan-sharpened images are depicted in Table 5 for both Landsat-8 and WorldView-2 satellite images. As it is obvious from no-reference based quantitative evaluation results of Landsat-8 OLI pan-sharpened products, Gram-Schmidt and PCA have fewer amounts of standard deviations, respectively. In WorldView-2 pan-sharpened products also Gram-Schmidt and PCA have the most entropy and the fewer amounts of standard deviations, respectively.

Discussion and Conclusion

In this paper, the pan-sharpened products obtained from PCA, IHS, and Gram-Schmidt transformation based methods are evaluated concerning the five specific spectral, spatial, and content-based measures. In addition to comparing the efficiencies of three different pan-sharpening methods, this study used satellite images with medium and high spatial resolutions those are taken from the same area to determine the impact of increasing the spatial resolution on the quantitative assessment results of pan-sharpened products.

Table 5: No-reference quantitative assessment of pan-sharpened images

Methods	Landsat-8 OLI		WorldView-2	
	Entropy	Standard Deviation	Entropy	Standard Deviation
PCA	4.9996	61.0933	7.3725	53.8688
IHS	**6.8784**	66.6764	6.6915	62.3462
Gram-Schmidt	5.1032	**56.8920**	**7.3749**	**51.2074**

In the spectral analysis, Gram-Schmidt has the least spectral deviations in both WorldView-2 and Landsat-8 OLI satellite images. The limitation on the number of spectral bands that can be processed in IHS pan-sharpening algorithm (only three spectral bands) is the main reason for more spectral deviations in IHS based pan-sharpened products. Moreover, replacing the first component of the multispectral image in the PCA transformation space with the panchromatic image causes the loss of spectral information of this component. Therefore, the PCA pan-sharpened products have some more spectral deviations than Gram- Schmidt and IHS results (Fig. 3).

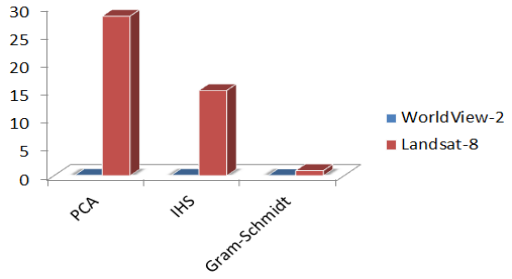


Fig. 3: Spectral quantitative assessment comparison based on PPD measure.

On the other hand, these results confirm the fact that differences in spatial resolution can't affect the Gram-Schmidt pan-sharpening results. However, as it is depicted in Fig. 3, the spectral deviations in the pan-sharpened products of WorldView-2 are less than Landsat-8 pan-sharpened images. The spectral deviations have been reduced for about 0.8658 in Gram-Schmidt, 15.0392 in IHS, and 28.2258 for PCA pan-sharpened products by increasing the spatial resolution of the data from 15 meters to half a meter. Therefore, increasing the spatial resolution can decrease the amounts of spectral deviations in the pan-sharpened images. Spatial analysis of the quantitative assessment results indicates that the most spatial correlations between pan-sharpening products and panchromatic image belong to the Gram-Schmidt pan-sharpening algorithm and PCA and HIS algorithms are in the next grades. According to the spatial analysis results, increasing the spatial resolution of the WorldView-2 pan-sharpened products regarding the Landsat-8 led to increase edge extraction quantitative evaluation results for about 24.16% for PCA, 4.69% for IHS and 12.48% for Gram-Schmidt pan-sharpening algorithms. Moreover,

high-pass filter quantitative assessment results also confirm the advantage of performing pan-sharpening algorithms on the high spatial resolution images. According to the results, most changes in the high-pass filter measure occurred on the PCA pan-sharpening method with 63.43% and then, for the IHS method with 19.96% (Fig. 4).

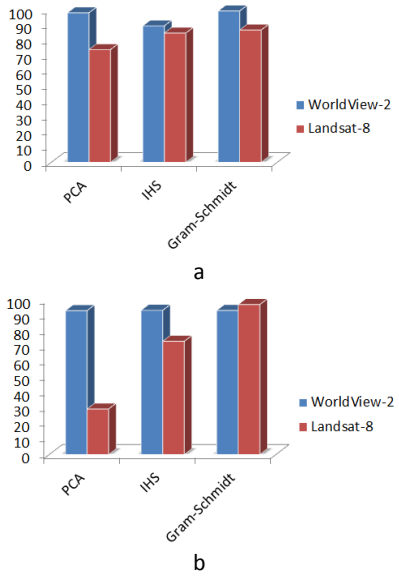


Fig. 4: Spatial quantitative assessment comparison based on a) Edge extraction metric, b) High-pass filter metric.

Content based quantitative assessment measures don't show a significant difference between the efficiencies of Gram-Schmidt, PCA and IHS pan-sharpening algorithms. However, as it is shown in Fig. 5 (a), the amount of entropy is increased with a higher spatial resolution of WordView-2 in PCA (for about 2.3729) and Gram-Schmidt (for about 2.2717) pan-sharpening method.

Increasing the spatial resolution almost doesn't have any impact on the IHS pan-sharpening results. Moreover, standard deviations of the PCA, IHS, and Gram-Schmidt pan-sharpened products of WorldView-2 reduced for about 7.2245, 4.3302, and 5.6846, respectively. As a general conclusion, it can be said that Gram-Schmidt pan-sharpening method has the best performance in both medium-scale and high-resolution data sets based on the spectral, spatial, and content quantitative evaluation results.

The IHS pan-sharpening method has better performance than the PCA method in Landsat-8 OLI data. But, by increasing the spatial resolution of the data, PCA generates pan-sharpened products with better spectral, spatial, and content based quantitative evaluation results. Therefore, it can be concluded that for medium-scale remote sensing data pan-sharpening, the IHS method can be a good choice. But, for high spatial resolution data such as WorldView-2, the PCA has better pan-sharpening results than the IHS.

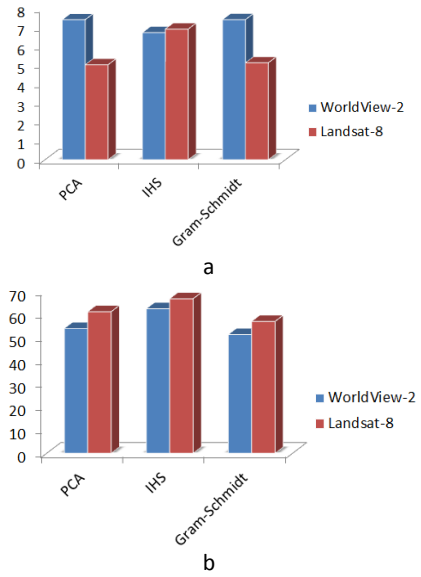


Fig. 5: Quantitative assessment comparison based on a) Entropy, b) Standard deviation

Since, the main objectives of this paper were investigating the efficiencies of transformation based pan-sharpening algorithms based on spectral, spatial and content quantitative assessment metrics, future researches can be conducted in utilizing wavelet algorithm as another transformation based pan-sharpening method. Moreover, applying the capabilities of the Spectral Angle Mapper (SAM) for spectral deviation analysis and Cross Entropy for spatial correlation measurement of the pan-sharpening products should be investigated. Confirming the obtained results in this paper about the well-known pan-sharpening algorithms can eliminate the labor intensive activities for selecting the optimum pan-sharpening algorithm in each image processing or pattern recognition application.

Author Contributions

A. Karami collected the data and applied the quantitative assessment metrics on the pan-sharpened products. F. Tabib Mahmoudi carried out the data analysis, interpreted the results and wrote the manuscript.

Conflict of Interest

The author declares that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

Aknowledgement

The authors gratefully acknowledges the supports provided by Geomatics department (Remote Sensing branch) of the Civil engineering from Shahid

Rajaee Teacher Training University, Tehran, Iran.

Abbreviations

IHS	Intensity-Hue-Saturation
PCA	Principal Component Analysis
GS	Gram-Schmidt
PPD	Per Pixel Deviation
OLI	Operational Land Imager
SAM	Spectral Angle Mapper

References

[1] G. Vivone, L. Alparone, J. Chanussot, M. Dalla Mura, A. Garzelli, G.A. Licciardi, R. Restaino, L.Wald, "A Critical Comparison Among Pansharpening Algorithms," IEEE Transactions on Geoscience and Remote Sensing, 53(5): 2565-2586, 2015.

[2] S. Yang, M. Wang, L. Jiao, "Fusion of Multispectral and Panchromatic Images Based on Support Value Transform and Adaptive Principal Component Analysis" Information Fusion, 13(3): 177-184, 2012.

[3] M. Deshmukh, U. Behosale, "Image Fusion and Image Quality Assessment of Fused Images" International Journal of Image Processing (IJIP), 4(5): 484-508, 2010.

[4] O.A. Agudelo-Medina, H. Dario Benitez-Restrepo, G. Vivone, A. Bovik, "Perceptual Quality Assessment of Pan-Sharpened Images," Remote Sens. 11(7): 1-19, 2019.

[5] A. Makarau, G. Palubinskas, P. Reinartz, "Analysis and selection of pan-sharpening assessment measures," Journal of Applied Remote Sensing, 6 (1): 1-20, 2012.

[6] C. Pohl, J.L. Van Jenderen, "Review article Multisensor image fusion in remote sensing: concepts, methods and applications," International Journal of Remote Sensing, 19(5): 823 -854, 2010.

[7] S. Aghapour Maleki, H. Ghassemian, "A critical review of quality assessment protocols in pan-sharpening," in Proc. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-4/W18, Joint Conferences of SMPR and GI Research: 13-18, 2019.

[8] G. Palubinskas, "Quality assessment of pan-sharpening methods," presented at 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada, 2014.

[9] P. Mhangara, W. Mapurisa, N. Mudau, "Comparison of Image Fusion Techniques Using Satellite Pour l'Observation de la Terre (SPOT) 6 Satellite Imagery," Applied Sciences, 10(5): 1-13, 2020.

[10] A.G. Sulaiman, W.H. Elashmawi, G.S. El-Tawel, "A Robust Pan-Sharpening Scheme for Improving Resolution of Satellite Images in the Domain of the Nonsubsampled Shearlet Transform," Sensing and Imaging, 21(3), 2019.

[11] S.M.A. Wady, Y. Bentoutou, A. Bengermikh, A. Bounoua, N. Taleb, "A new HIS and wavelet based pansharpening algorithm for high spatial resolution satellite imagery," Advances in space research, 66(7): 1507-1521, 2020.

[12] Y. Choi, E. Sharifahmadian, Sh. Latifi, "Quality Assessment of Image Fusion Methods in Transform Domain," International Journal on Information Theory (IJIT), 3(1): 7-18, 2014.

[13] S. R. Dammavalam, S. Maddala, K. Prasad MHM, "Quality assessment of pixel-level image fusion using fuzzy logic," International Journal on Soft Computing (IJSC), 3(1): 13-15, 2012.

[14] H.B. Mitchell, "Image fusion, theories, techniques and applications," Springer-Verlag Berlin Heidelberg, Germany, 2010.

[15] T. Stathaki, "Image fusion, algorithms and applications," Academic Press is an imprint of Elsevier, Britain, 2008.

[16] M.R. Metwalli, A.H. Nasr, O.S. Farag Allah, S. El-Rabaie, "Image Fusion Based on Principal Component Analysis and High-Pass Filter," International Conference on Computer Engineering & Systems, Cairo: 63-70, 2009.

[17] V.P.S. Naidu, and J.R. Raol, "Pixel-level Image Fusion using Wavelets and Principal Component Analysis," Defence Science Journal, Vol. 58 (3), pp. 338-352, 2008.

[18] T.M. Tu, S.C. Su, H.C. Shyu, P.S. Huang, "A New Look at IHS- Like Image Fusion Method," Information Fusion, 2(2001): 177-186, 2001.

[19] C. Yang, Q. Zhan, H. Liu, R. Ma, "An IHS-Based Pan-Sharpening Method for Spectral Fidelity Improvement Using Ripplet Transform and Compressed Sensing," Sensors, 18(11): 1-20, 2018.

[20] F. Samadzadegan, F. Tabib Mahmoudi, "Data Fusion in remote sensing concepts and techniques," University of Tehran Press 3316, 2nd Edition, 2014.

Biographies

Fatemeh Tabib Mahmoudi received her B.Sc. degree in civil engineering the branch of Geomatics (Surveying and mapping) from Khajeh Nasiredin Toosi University, Tehran, Iran, in 2004. She received his M.Sc. and PhD degrees in Photogrammetry from Tehran University, Tehran, Iran, in 2009 and 2014, respectively. Since 2016, she has been working as an assistant professor in the Geomatics department of the faculty of Civil Engineering, Shahid Rajaee Teacher Training University. She has some publications in the field of remote sensing data analysis, pattern recognition and data fusion.

Adel Karami received his B.Sc. degree in civil engineering the branch of Geomatics (Surveying and mapping) from Tabriz University, Tabriz, Iran, in 2019. Since 2020, he is the M.Sc. student of remote sensing in Geomatics department of the faculty of Civil Engineering, Shahid Rajaee Teacher Training University.

How to cite this paper:
F. Tabib Mahmoudi, A. Karami "Quantitative Assessment of Transformation Based Satellite Image Pan-sharpening Algorithms," Journal of Electrical and Computer Engineering Innovations, 8(2): 161-168, 2020.

DOI: 10.22061/JECEI.2020.7191.367

URL: http://jecei.sru.ac.ir/article_1458.html

Journal of Electrical and Computer Engineering Innovations
(JECEI)
Journal homepage: http://www.jecei.sru.ac.ir

Research paper

An Energy Efficient Fault Tolerance Technique Based on Load Balancing Algorithm for High-Performance Computing in Cloud Computing

H. Jahanpour, H. Barati*, A. Mehranzadeh

Department of Computer Engineering, Dezful Branch, Islamic Azad University, Dezful, Iran.

Article Info

Article History:
Received 12 November 2019
Reviewed 27 December 2019
Revised 02 February 2020
Accepted 04 May 2020

Keywords:
Cloud computing
Fault tolerance
High-Performance computing
Virtual machines
Load balancing

*Corresponding Author's Email Address:
hbarati@iaud.ac.ir

Abstract

Background and Objectives: Cloud Computing has brought a new dimension to the IT world. The technology of cloud computing allows employing a large number of Virtual Machines to run intensive applications. Each failure in running applications fails system operations. To solve the problem, it is required to restart the systems.

Methods: In this paper, to predict and avoid failure in HPC systems, a method of fault tolerance to High-Performance Computing systems (HPC) in the cloud is called Daemon-COA-MMT (DCM), has been proposed. In the proposed method, the Daemon Fault Tolerance technique has been enhanced, and COA-MMT has been utilized for load balancing. The method consists of four modules, which are used to determine the host state. When the system is in the alarm state, the current host may face failure. Then the most optimal host for migration is selected, and process-level migration is performed. The method causes decreased migration overheads, decreased system performance speed, optimal use of underutilized hosts instead of leasing new hosts, appropriate load balancing, equal use of hardware resources by all hosts, focusing on QoS and SLA, and the significant decrease of energy consumption.

Results: The simulation results revealed that in terms of parameters, the proposed method declines average job makespan, average response time, and average task execution cost by 18.06%, 35.68%, and 24.6%, respectively. The proposed fault tolerance algorithm has improved energy consumption by 30% and decreased the HPC systems' failure rate.

Conclusion: In this study, the Daemon Fault Tolerance technique has been enhanced, and COA-MMT has been utilized for load balancing in high performance computing in the cloud computing.

Introduction

Cloud computing is the greatest revolution in the computing world, so that significant organizations and companies have changed their traditional data processing systems to cloud service to store a large amount of data [1]. Cloud computing advantages are running computation-intensive applications, decreased time of applying the hardware, and cost [2]. It reduces the time of applying the hardware and cost. There are two critical roles in cloud computing: cloud service providers and users [3]. The providers such as Amazon and Bare Metal Cloud offer Virtual Machines (VMs), the hardware, etc. to their clients in return for the subscription. Based on the services provided by them, clouds are divided into four categories: Infrastructure as

a Service (IaaS), Software as a Service (SaaS), Platform as a Service (PaaS), and Hardware as a Service (HaaS) [4, 5]. HaaS focuses on the hardware.

The service can be leased for research, massive information, and configuration of HPC systems [4]. HPC is a branch of software science that causes great scientific and computing jobs so rapidly and less costly by integrating the computing power of many small and medium computers [6]. HPC systems can process a large volume of data and analyze the results so rapidly. Using HaaS, running conventional computation-intensive applications on HPC systems in the cloud will be possible [7]. Despite different advantages such as fastness, resource provisioning, cost reduction, multitenant services, etc., cloud computing faces various challenges, including load balancing, security, reliability, possession, green technology, backing up data, and transferring data [8]. Some of the most critical cloud computing challenges are reliability and resource availability, especially at the HaaS level [9].

A system will be called fault tolerance if it fulfills its determined duties properly, even in the presence of software and hardware failures [10]. In fault-tolerance systems, the system's restarting is refrained to decrease operational costs and energy consumption [11]. The importance of fault tolerance is to develop the availability of resources, reliability of cloud services, and running applications. To minimize the effects of a failure on the system and provide accurate and successful running of applications, failures should be predicted and managed [12]. If fault tolerance is not provided, the system will incur irreparable damage [13]. Therefore, fault tolerance is an essential feature of cloud computing systems, especially HPC systems, since it results in shorter running times in the presence of failure. Also, load balancing is one of the main challenges of cloud computing, which divides workload evenly between hosts to satisfy users and increase the rate of resource consumption [14]. Load balancing aims to minimize energy consumption and reduce carbon dioxide emissions in cloud computing [15]. Decreased energy consumption in cloud computing systems leads to less carbon dioxide in cloud infrastructures, which causes less warming and pollution of the environment. Less energy consumption and carbon dioxide emission are essential criteria for energy-efficient load balancing in cloud computing, which causes green computing [16]. To provide energy-efficient fault tolerance, increase Quality of Service (QoS), cause effective use of resources, lessen violation of Service Level Agreement (SLA), reduce response time, and accurately examine the system's state. Then the failure is predicted and refrained through effective use of resources. This method causes energy-efficient fault tolerance and proper load balancing among hosts.

In the proposed method, the Daemon Fault Tolerance technique has been enhanced, and COA-MMT has been utilized for load balancing. The proposed method consists of four modules: node monitoring with Lm sensors module, rule-based predictor module, migration policy module based on COA-MMT, and controller module of DCM. The method causes decreased migration overheads, decreased system performance speed, optimal use of underutilized hosts instead of leasing new hosts, appropriate load balancing, equal use of hardware resources by all hosts, focusing on QoS and SLA, and a significant decrease of energy consumption.

The paper is organized as follows: Next Section includes related work in fault tolerance and load balancing. In Next Section, the proposed method, DCM, and its modules for fault tolerance in HPC systems are discussed in detail. Next Section consists of the simulation and evaluation of the method. Finally, Section presents the conclusions.

Related Work

This part investigates specific algorithms in load balancing and fault tolerance in cloud computing and individually represent their advantages and disadvantages.

Pan et al. [17] represented Particle Swarm Optimization (PSO) algorithm, an evolutionary computing method, originated from particles' social and natural behavior. Particles possess state and speed and move in a multidimensional search space. Each particle determines its speed based on its own best state and the state of the best particle in the society, which reduces response time [18-19].

Huang et al. [20] suggested the Genetic Algorithm. In this algorithm, the gene cost is first calculated through the current scheduling solution's ratio to the best scheduling solution. Then based on the gene cost, a scheduling strategy is decided. Finally, the least costly solution, which is similar to the final scheduling solution, is selected. The standard genetic algorithm guarantees the load balancing of the system more effectively compared with other methods. The rate of loading fluctuation of VMs plays an essential role in load balancing.

Abdullah et al. [21] suggested that Bat Algorithm provides load balancing. In this method, first, each bat receives a primary value. The speed and location of each bat are randomly determined in a d-dimensional space. Each bat's fitness function is calculated, and the best state for the bat is determined based on the least value of the function. This method provides optimal utilization of all resources and is more efficient. Its convergence speed is superior to that of PSO. However, an increased number of requests causes a longer response time. Also,

users wait a long time to receive service. The result will be decreased QoS and users' satisfaction and increased violation of SLA and system costs.

Ghafari et al. [22] represented the Artificial Bee colony Algorithm-Minimal Migration Time (Bee-MMT) to provide load balancing. In this method, first, the over-utilized host is determined. Then, one or more VMs are determined to migrate to underutilized hosts. Then VMs migrate from over-utilized hosts to new hosts. On migrating, the previous host switches to sleep mode. In this algorithm, the violation of SLA is used as an essential metric to satisfy QoS. The method provides better response time than conventional methods and reduces energy consumption in cloud computing infrastructure. Also, BEE-MMT causes decreased carbon dioxide and the appropriate efficiency of the resources of the system.

In [23], Daemon's fault tolerance is proposed. This algorithm is based on the methods of predicting failures. This algorithm has four modules with specific duties such as node monitoring module with Linux monitoring sensors, rule-based fault predictor module, migration policy module, and controller module. The first module is the node monitoring module Lm sensors. The monitoring node is an open-source using Lm sensors to monitor the accuracy of the computer's tasks. Modern CPUs are made of sensors used for monitoring CPU temperature, fan speed, memories, number of user's requests, and other parameters. Rule-based prediction is the second stage. At this stage, the failure is predicted based on the history of failures and the system's maximum workload. The predictor module inputs consist of four parameters, to which specific weight values are assigned to calculate the state of the operating system. In the third stage, the migration policy is implemented. The purpose of the policy is to execute computation-intensive entirely with minimum energy consumption. At the fourth stage, the controller module is implemented. As failure is predicted, FTDaemon calls for the module. Occurring failure requires the system to lease an additional node. On migrating from an unhealthy node to a newly leased node, the unhealthy node is abandoned. When the host is not operating at a critical state, there is no need to keep an additional node so that additional nodes' cost and energy are nearly zero. In FTDaemon, when a host is predicted to fail, the system manager will lease a new host from the service provider. This is a main weakness of the method since other hosts, which may be underutilized, will not be considered by the manager, so load balancing is not established.

In [24], reactive fault tolerance is suggested. In this method, while HPC systems are running, they send their results to checkpoints. In case of any failure, the system restarts from the point before the failure. In this method, due to increased system components, the system may not be able to restart repeatedly. The technique leads to decreased energy consumption. Also, the method does not suit the systems needing overuse of VMs or clusters because failures lead to a significant decrease in availability.

In [25], Power-Check fault tolerance has been suggested, which increases the monitoring level in HPC systems using specific intelligent data. The method causes decreased CPU use, lower system performance, and optimized energy consumption.

Yakhchi et al. [26] suggested Cuckoo Optimization Algorithm-Minimum Migration Time (COA-MMT) algorithm to provide load balancing. This algorithm is based on the life of cuckoos. COA-MMT has three steps for load balancing and power consumption management: At the first stage, an over-utilized host is detected. To do this, some hosts are selected randomly and clustered. Using profit function and according to equation (1), the profit value of habitat or cluster is determined. Applying equation (2), for each host, the eggs are laid in a specific range called Egg Laying Radius (ELR) [16]. The host with the most CPU utilization is selected as the overused host.

$$profit = f_p(habitat) = f_p(x_1.x_2.\ldots.x_{Nvar}) \quad (1)$$

In equation (1), f_p denotes the profit function.

$$ELR = \alpha * \frac{Number\ of\ current\ cuckoos\ eggs}{Total\ number\ of\ eggs} * (var_{hi} - var_{low}) \quad (2)$$

In equation (2), α is an integer, supposed to handle the maximum value of ELR. var_{hi} and var_{low} stand for upper bound and lower bound, respectively, which are used for defining ELR. At the second stage, an under-loaded host is detected. The host experiencing the minimum CPU utilization is selected as the host with the least loading value. At the third stage, selection policy is implemented, and one or more VMs are selected to migrate to the host with minimum CPU utilization. Minimal Migration Policy Time (MMT) selects the VMs needing less time to migrate to other hosts. Migration time is calculated by equation (3).

$$v \in V_j \mid \forall\, a \in V_j, \frac{RAM_u(v)}{NET_j} \leq \frac{RAM_u(a)}{NET_j} \quad (3)$$

In equation (3) V_j is a set of VMs currently allocated to host j. NET_j denotes the spare network bandwidth available for the host j; and $RAM_u(a)$ is the amount of RAM currently utilized by the VM a. This method decreased the violation of SLA compared with Bee-MMT. Using this method leads to increased QoS and satisfaction of users.

Tamilvizhi and Parvathavarthini in [27] proposed a concept of fault management with the emphasis on the

hardware and network faults handling. This proposed work introduces an innovative perspective on adopting a fault-tolerant mechanism to avoid network congestion and health monitoring for fault detection with migration techniques to handle faults adaptively. This work's primary goal is to develop an effective cloud architecture that could tolerate fault occurrences beforehand or after hand and then suggest appropriate solutions to maintain data traffic and the system's availability, thus making it more reliable and flexible.

Neelima and Reddy in [28] proposed a load balancing task scheduling algorithm in the cloud using the Adaptive Dragonfly algorithm (ADA), which provides minimum time and cost while balancing the load. In this method, to attain better performance, a multi-objective function is developed based on three parameters: completion time, processing costs, and load. Based on the multi-objective function, we assign a task to VM. The proposed methodology's main objective is to assign the task to VM using ADA, which minimizes the total execution time and cost while balancing the load.

Durga Devi et al. [29] proposed a dynamic load balancing in a heterogeneous environment by Modified Adaptive Neuro-Fuzzy Inference System (MANFIS). Parameters of MANFIS are optimized by introducing Fire-fly Algorithm. In this method, the adopted Modified Adaptive Neuro-Fuzzy Inference System (MANFIS) for VM load balancing is based on the CPU utilization and turnaround time. Also adopted Enhanced Elliptic Curve Cryptography to provide security between cloud users and cloud servers. There are two key implication of proposed methodology. First, is to optimize load balancing based on CPU utilization and Turnaround time. Second, is to provide data security using Enhanced Elliptic Curve Cryptography.

Kong et al. [30] proposed a fast heuristic algorithm based on the zero imbalance approach as a new concept in the heterogeneous environment. This approach focuses on minimizing the completion time difference among heterogeneous VMs without priority methods and complex scheduling decision to the particular cloud configuration. This mechanism consists of combining load balancing and task allocation. To achieve this mechanism, this algorithm collects each task's size, the processing speed of each VM, the bandwidth of each VM, the number of VMs and tasks, as information to implement load balancing and task allocation in the balancing phase. Moreover, the assignment of tasks is performed on any VMs under the control of modified optimal completion time. The proposed algorithm identifies the suitable VMs for the appropriate unassigned tasks based on earliest finish time in the task allocation phase. Table 1 shows a comparison between the mentioned algorithms.

Proposed Method

As shown in the related work section, FTDaemon is not perfect, which results in increasing migration overheads, increasing system cost, increasing energy consumption, decreasing system performance speed, ignoring underutilized hosts, inappropriate load balancing, unequal use of hardware resources by some hosts, ignoring QoS, and violating SLA. An energy-efficient fault tolerance approach has been suggested to predict and avoid failure occurrence in HPC systems. The proposed algorithm is called Daemon-COA-MMT (DCM). The method causes decreased migration overheads, decreased system performance speed, optimal use of underutilized hosts instead of leasing new hosts, appropriate load balancing, equal use of hardware resources by all hosts, focusing on QoS and SLA, and a significant decrease of energy consumption. Our method employs four modules. To predict and prevent failures, the proposed method utilizes these parameters: CPU temperature, CPU utilization, number of users' requests, voltage, and fan speed parameters. The architecture of the method consisted of some modules, is illustrated in Fig. 1.

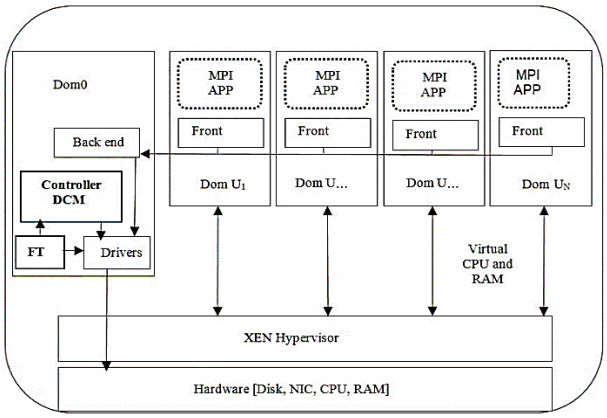


Fig. 1: The architecture of the proposed method.

The proposed method consists of four modules:

- Node monitoring with Lm sensors module
- Rule-based predictor module
- Migration policy module based on COA-MMT
- Controller module of DCM

The modules are described as follows:

A. Host Monitoring Modules in Proposed Method

In the proposed method, Lm sensors are used since most HPC systems run Linux, and Lm sensors utilize the Linux operating system. Lm sensors cause the development of the DCM method, which could easily be deployed on HPC systems in clouds. HPC systems, possessing more than 100000 CPUs, impose massive overhead on the networks and HPC systems. Therefore,

the method should check the parameters periodically to reduce monitoring overhead.

The information collected at intervals of 600 seconds, which are changeable, is sent to the host. Whenever the monitoring parameters inside the Lm sensors exceed the maximum value, the alarm will be triggered, which indicates that the failure is likely to occur.

B. Rule-Based Prediction Modules in Proposed Method

DCM Fault Tolerance is executed on each node in the user's space, and the failure is predicted based on the history of failure, maximum operating values, and information obtained from the system.

When the monitoring node with Lm sensors indicates a failure, the rule-based predictor module runs. The rule-based predictor module inputs are five parameters: temperature T, voltage V, fan speed F, CPU utilization C, and several user's requests R from the host.

The reason for using number of user's requests is that the second module investigates hardware resources and the cause of creating workload for the host's hardware resources. To calculate the respective host's actual state, specific fixed weight values are assigned to the host. The values are 1, 1.5, 2, 2.5, and 3 for the very good, good, normal, alarm, and critical areas, respectively (Table 2).

Table 1: Comparison and summary of previous methods

Ref	Approach	Advantage	Disadvantage
[17]	An improved particle algorithm to achieve resource load balancing optimization in the cloud environment	Improve Resource utilization, Good Performance	It is valid for equal-sized population
[20]	a Genetic Algorithm based resource management algorithm for allocating cloud-based virtual machines on physical machines	Obtained an optimized distribution strategy	High computational overhead
[21]	The comparison of load balancing techniques and BAT algorithm techniques are described	provides optimal utilization of all resources	Increased number of requests causes a longer response time
[22]	An algorithm to detect over utilized hosts and then migrate VMs based on artificial bee colony algorithm (ABC)	Greater power consumption saving, Decreasing the CO2 emission and operational cost	High complexity for selecting best overloaded host, No prediction for future workload of hosts
[23]	Energy efficient fault tolerance for HPC in the cloud that develop a generic FT algorithm for HPC systems in the cloud.	Reduced the energy consumption of computation-intensive applications	Low accuracy of failure prediction mechanism that is unsuitable for HPC workload.
[24]	Fault Tolerance (FT) approach to HPC systems in the cloud to reduce the wall clock execution time in the presence of faults	Improved the execution time, reduce energy consumption	Does not suit the systems needing overuse of VMs or clusters
[25]	A power-aware check pointing framework Power-Check to address the problem of marginal energy benefits	Reduction in the amount of energy consumed, improving the check pointing performance	Job partitioning however not considered in this approach.
[26]	An approach based on Cuckoo Optimization Algorithm (COA) to detect over-utilized hosts.	Reduced the power consumption	May be cause SLA violation
[27]	Adopting a fault tolerant mechanism to avoid network congestion and health monitoring for fault detection with migration technique	Reduced energy consumption and cost overhead	This method no worries about how to cover the error
[28]	A load balancing task scheduling algorithm in cloud using Adaptive Dragonfly algorithm (ADA)	Well-balanced load across virtual machines	High computational overhead
[29]	Dynamic load balancing in a heterogeneous environment is handled by Modified Adaptive Neuro Fuzzy Inference System (MANFIS)	Improving the turnaround time and maximizing the CPU utilization	High communication overhead
[30]	A fast heuristic algorithm based on the zero imbalance approach, as a new concept in the heterogeneous environment	strikes the balance between the requirements of cloud users and providers	Ignoring power consumption in the data center and live VM migration.

Table 2: Weight of parameters based on measured

CPU Utilization	Number of Users' Requests from Host	Fan Speed	Voltage	Temperature	Weight of Parameter
0-16	0-50	0-500	0.94-0.85	0-15	1
16-32	50-100	500-1000	0.94-1.03	15-30	1.5
32-48	100-150	1000-1500	1.03-1.12	30-45	2
48-64	150-200	1500-2000	1.12-1.21	45-60	2.5
64-80	200-250	2000-2500	1.21-1. 30	60-75	3

As shown in Table 2, the host is at an excellent state when parameters are as follows: temperature 0-15, voltage 0.85-0.94, fan speed 0-500, CPU utilization 0-16, and the number of requests 0-50. The weight of each parameter is 1. The rule-based predictor module is shown in Fig. 2.

As shown in Fig. 2, the main parameters $(T_i, F_i, V_i, C_i, R_i)$ are inserted into a calculating module and a_c values are obtained from equation (5). The result is compared with the threshold, and the output of the rule-based predictor module is obtained. The threshold is calculated based on system log, constructive information, current sensor values, and CPU utilization values.

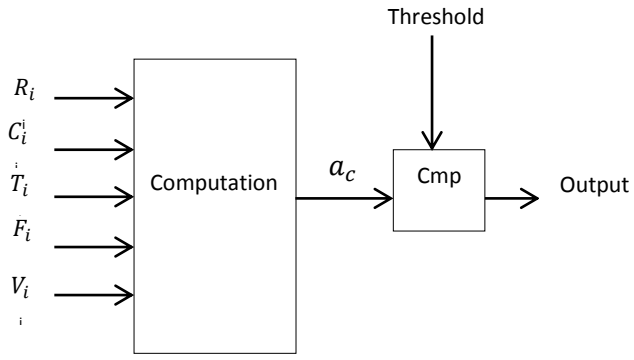


Fig. 2: Rule based predictor module.

In (4), K_c constant is employed to improve the accuracy of prediction. As K_c is negligible, we set $K_c = 0$. On calculating a_c based on determining ranges in (5), the host state is determined based on ac at the current state.

$$a_c = \prod_{i=1}^{n} T_i * F_i * V_i * C_i * R_i + K_c \quad (4)$$

As shown in (6), five states can be assigned to each host. The categorization is based on these five parameters: temperature (T), voltage (V), fan speed (F), CPU utilization (C), and the number of user's requests (R) from the host. These five areas are used to improve the accuracy of prediction and detect suspicious hosts immediately. The method of determining the state of hosts is shown in Table 3.

$$a_c = \begin{cases} Critical\ State & 117.18 \le |a_c| \le 243 \\ Alarm\ State & 40 \le |a_c| \le 97.66 \\ Normal\ State & 10.12 \le |a_c| \le 32 \\ Good\ State & 1.5 \le |a_c| \le 7.59 \\ Very\ Good\ State & |a_c| = 1 \end{cases} \quad (5)$$

C. Migration Policy Based on Proposed Method

When an alarm is triggered, the migration policy is activated.

It is not needed to lease a new host to eliminate the alarm state since the third module examines all hosts to find underutilized hosts.

The purpose of the method's migration policy is to complete computation-intensive computation with minimum energy consumption. In the DCM algorithm, when a failure is predicted, COA-MMT load balancing is executed.

On investigating the load balancing area, we chose the COA-MMT load balancing algorithm for the third module due to its rapid and exact detection of optimal point, providing appropriate load balancing and SLA, increasing QoS, and containing MMT policy. COA-MMT technique is executed in three steps to provide load balancing in the system.

In the third module of the proposed method, the COA-MMT load-balancing algorithm is implemented in two steps to establish load balancing and manage power utilization. According to the method, the host monitoring and rule-based predictor modules are determined based on the over-utilized host's hardware parameters. Hence, the COA-MMT load balancing algorithm does not need to search for the over-utilized host, which causes overheads and increased energy consumption. Migration policy based on COA-MMT optimization includes two steps: detecting the under loaded host and selection policy.

Table 3: Calculating a_c based on the parameters of the proposed method

State	Amounts aC					Threshold
Critical	3*2.5*2.5*2.5*2.5 =117.18	3*3*2.5*2.5*2.5 =140.62	3*3*3*2.5*2.5 =168.75	3*3*3*3*2.5 =202.50	3*3*3*3*3 =243	117.18
Alarm	2*2*2*2*2.5 =40	2*2*2*2.5*2.5 =50	2*2*2.5*2.5*2.5 =62.50	2*2.5*2.5*2.5*2. 5 = 78.12	2.5*2.5*2.5*2.5* 2.5 = 97.66	40
Normal	1.5*1.5*1.5*1.5*2 =10.12	1.5*1.5*1.5*2*2 =13.50	1.5*1.5*2*2*2 =18	1.5*2*2*2*2 =24	2*2*2*2*2 =32	10.12
Good	1*1*1*1*1.5 =1.5	1*1*1*1.5*1.5 =2.25	1*1*1.5*1.5*1.5 =3.37	1*1.5*1.5*1.5*1. 5 = 5.06	1.5*1.5*1.5*1.5* 1.5 = 7.59	1.5
Very Good	1*1*1*1*1 =1	_	_	_	_	1

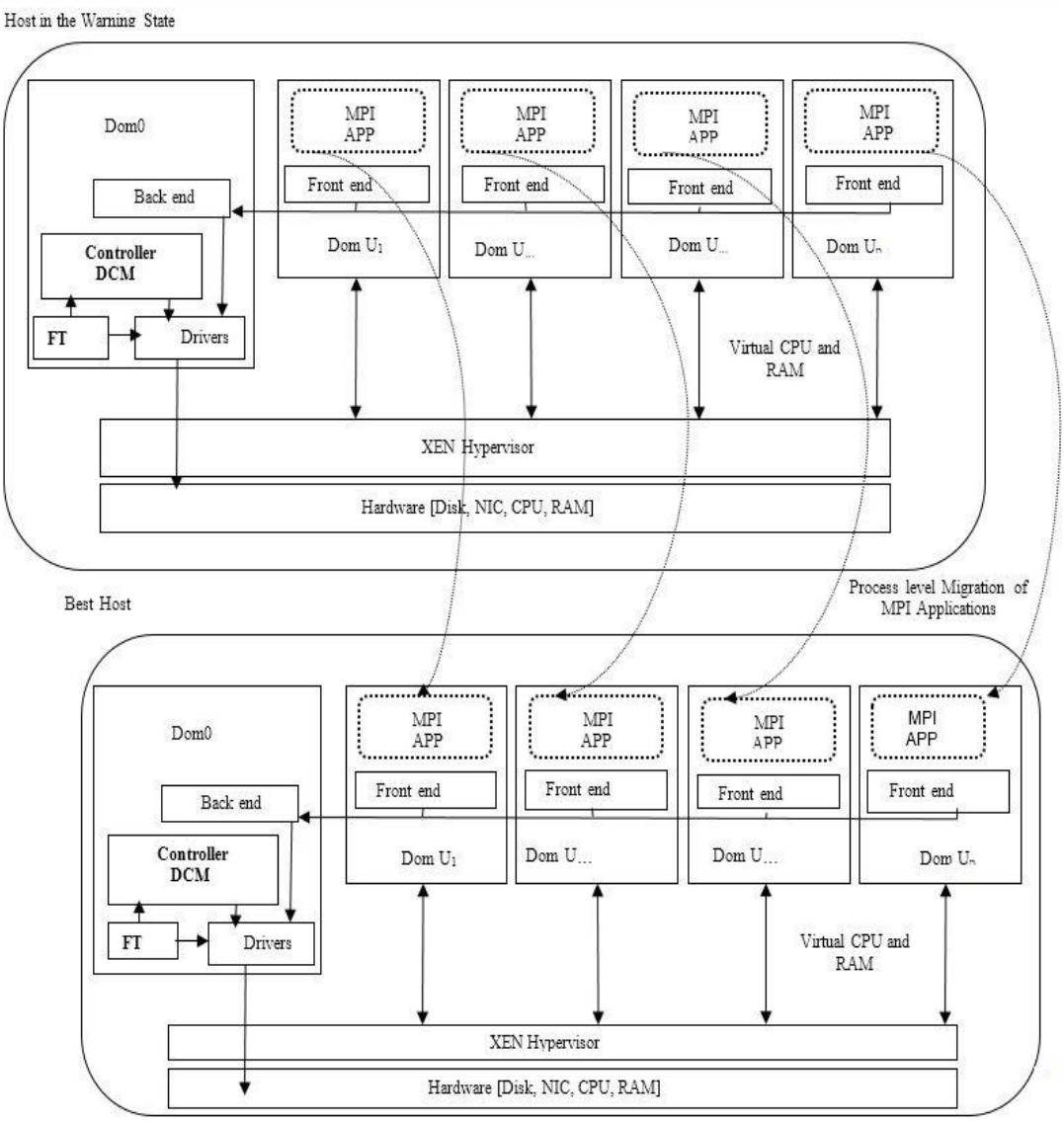


Fig. 3: Process level migration from the host, in warning state to the proper host.

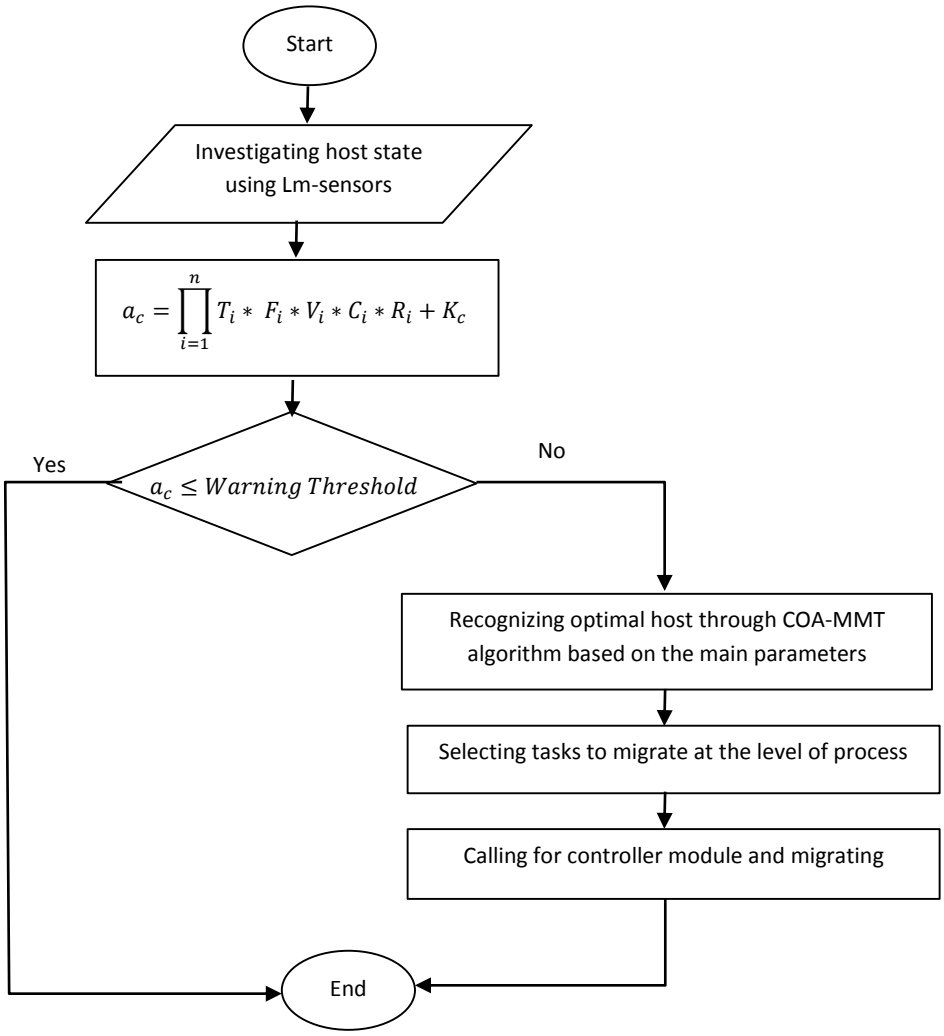


Fig. 4: Flowchart of the proposed method.

D. Controller Module in Proposed Method

The controller module is responsible for the implementation of three introductory modules. It is installed on all nodes. On predicting a failure, the controller module is called for, and the following steps are taken:

- Several VMs installed on a host, which is in the alarm state, request some information about their current host from the information center.
- The ID of programs running on the unhealthy VMs installed on the current host is obtained through the information center.
- An appropriate host, determined by COA-MMT, is selected
- The process level migration from a host in an alarm state to a host in the proper state is performed.
- The details of the running VMs on a proper host is published to the head host.

In Fig. 3, the process level migration from a host, which is in an alarm state to a proper host, is shown. Once the method detects a host, which is in an alarm state, the host is selected based on COA-MMT, and the controller module performs process level migration to a proper host.

The flowchart of the proposed method is illustrated in Fig. 4. First, the state of the host is examined by Lm sensors. Second, five main parameters are selected, and a_c is calculated through equation (5). The result is compared with the warning threshold. If a_c is lower than the warning threshold, the host state will be proper, and other steps are redundant. Otherwise, using the COA-MMT algorithm, the proper host is selected. The tasks are selected on the host, which is at the alarm state. Finally, process level migration is performed. Algorithm 1 shows the pseudo-code for proposed method.

Simulation and Result

Using Cloudsim 3.0, the proposed method, DCM, has been simulated. The proposed method's efficiency has been evaluated in scenarios A, B, and C compared with Power-Check [15] and Tamilvizhi et al. method [27]. In scenario A, five users with five brokers, and two data centers have been created. The first data center contains three hosts, while the second data center contains two hosts. Ten VMs are also created using the Time-Shared policy, each with 512 MB and one CPU managed by Xen, as Virtual Machine Manager (VMM), on Linux operating system. The host's memory is 2048 MB, with a storage capacity of 1,000,000 MB and a bandwidth of 10,000 Mb/sec. The number of submitted tasks (cloudlets) ranges between 10 and 100, each with 600 MB files.

In scenario B, we set 10 cloud users with ten brokers and five data centers. Each data center contains three hosts, making a total of 15 hosts. A total of 25 VMs are also created using the Time-Shared policy, each with 512 BM and one CPU managed by Xen, as VMM, on Linux operating system.

The host memory is 2048 MB, with a storage capacity of 1,000,000 MB and a bandwidth of 10,000 Mb/sec. Moreover, the number of submitted tasks ranges between 50 and 500, each with 1000 MB files.

In scenario C, fifteen cloud users with fifteen brokers and eight data centers have been created. Each data center contains three hosts, making a total of 24 hosts. 30 VMs are also created using the Time-Shared policy, each with 512 BM and one CPU managed by Xen, as VMM, on Linux operating system. The host's memory is 2048 MB, with a storage capacity of 1,000,000 MB and a bandwidth of 10,000 Mb/sec. Also, the number of submitted tasks ranges between 500 and 1000, each with a file size of 1400 MB. To improve accuracy, simulation is performed ten times in a row.

Table 4 illustrates the conditions and parameters of the simulation.

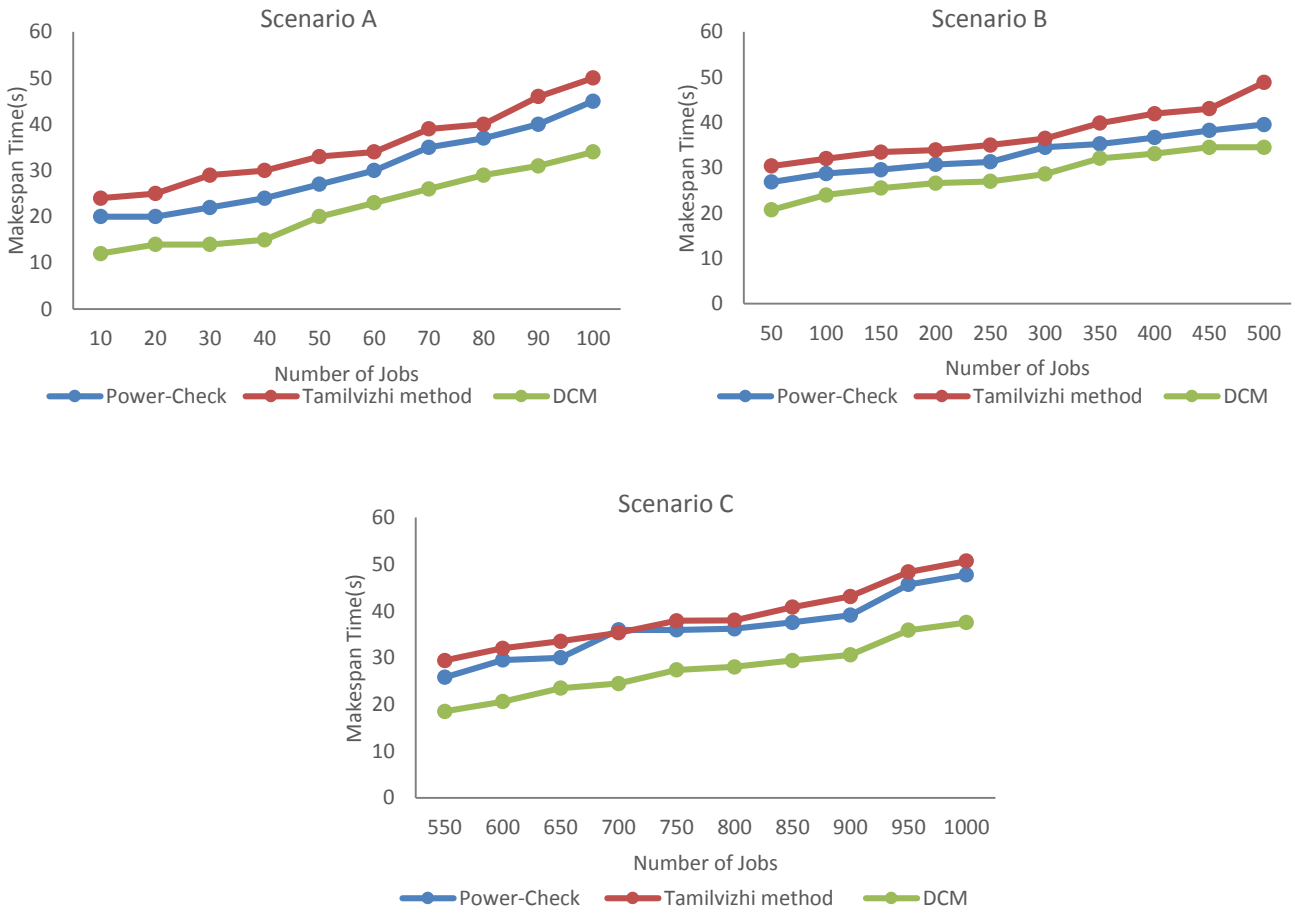


Fig. 5: Average job makespan. a.senario A's makespan time. b.senario B's makespan time. c. senario C's makespan time.

Table 4: Conditions and simulation parameters

Scenarios	A	B	C
Cloud users	5	10	15
Brokers	5	10	15
Data centers	2	5	8
Virtual Machines	10	25	30
Bandwidth (Mb/sec)	10000	10000	10000
Total Host	5	15	24
Number of Tasks	10 - 100	50 - 500	550 - 1000
Storage Capacity	1000000	1000000	1000000
host memory(MB)	2048	2048	2048

The average job makespan, average response time, failure rate, energy consumption, and average task execution costs are presented compared to two other algorithms.

The interval between request and completion of the request is called job makespan. Figure 5 shows the average job makespan of the proposed method in scenario A, B, and C compared with Power-Check and Tamilvizhi method.

Algorithm 1: Proposed method

Initialization: (n: number of host, T: temperature, V: voltage, F: fan speed, C: CPU utilization, R: several user's requests)
for (i=1 ; i<= n ;i++)
Using Lm sensors, the $host_i$ state mode is obtained $(T_i, F_i, V_i, C_i, R_i)$;
Determine weight of parameter for $(T_i, F_i, V_i, C_i, R_i)$;
Calculating ac values for $host_i$;
Determine the state of $host_i$;
if ($a_c > Threshold$)
Recognizing optimal host through COA-MMT algorithm;
Selecting tasks to migrate at the level of process
Calling for controller module and migrating
The details of the running VMs on a proper host is published to the head host.
end if
end for
end

As shown in Fig. 5, in scenario A, the average job makespan has improved compared with other methods. Also, an increased number of tasks results in improving the method. In scenario B, when the number of tasks changes from 50 to 500, which is more than that of scenario A, the average job makespan of the method is less compared with other methods. Also, in scenario C, when the number of tasks changes from 550 to 1000, the average job makespan improves substantially. Decreased average job makespan shows that the method's task execution time is lower compared with other methods.

The most optimal host is chosen for migration in the proposed method, and load balancing is established. Thereby, its average job makespan is decreased by 9.50% and 18.06%, respectively, compared with the Tamilvizhi method and Power-Check.

In the proposed method, first the important information of the nodes is collected using sensors and then the status of the nodes is checked for fault by using the prediction module. If a node is at faulty, jobs will be properly transferred to the appropriate machines by migration policy agents. In this way, a proper load balance will be created on the proposed method and makespan is reduced.

Fig. 6 shows the failure rate (FR), calculated by equation (6). Here, FR is calculated concerning the total failure of the system.

In the proposed method, by using the module to predict the status of the node and check the status of the node in terms of fault, an attempt is made to prevent fault in machines. Also, by performing the migration operation properly, a stable situation is provided to prevent fault. As shown in Fig. 6, increased workload and number of tasks cause the failure rate to decrease in HPC systems.

In the failure rate of the DCM method is decreased compared with the Tamilvizhi method and Power-Check by 21.03% and 10.21%, respectively.

Therefore, there is a relationship between average job makespan, failure rate, and reliability in HPC systems since the decrease of average job makespan leads to increased failure rate and reliability.

$$FR = \frac{1}{MTTF} \quad (6)$$

where FR is failure rate and MTTF is min time to failure. Equation (6) derives the rate of failure of our method concerning the system's total failure.

The interval between request and the first response is called response time.

The proposed method's response time compared with

the Tamilvizhi method and Power-Check in three considered scenarios is shown in Fig. 7.

As shown in Fig. 7, the method results in proper load balancing between all hosts. As mentioned, in the proposed method, by applying the appropriate migration method, a good load balancing is created.

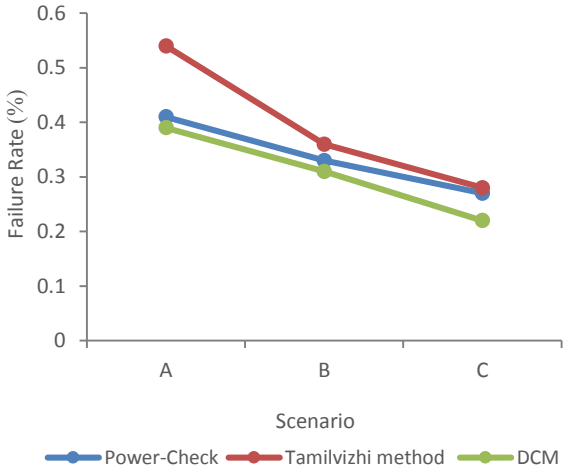


Fig. 6: Investigating the rate of failure in three scenarios.

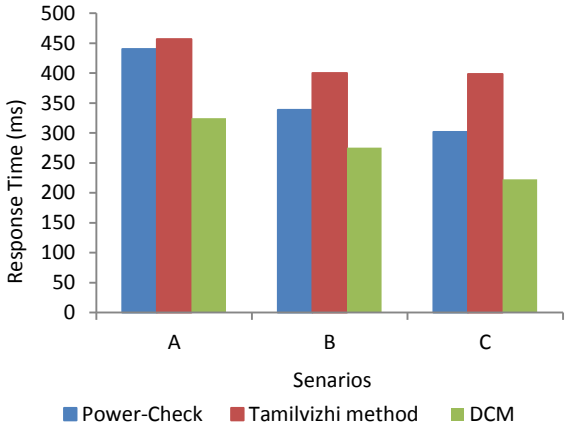


Fig. 7: Average Response Time.

Balancing the load between the machines makes the jobs on the machines faster. Also, the productivity of the CPUs of the system is increased. Therefore, the average response time is decreased compared with the Tamilvizhi method and Power-Check by 45.83% and 35.68%, respectively.

The costs that the service provider incur to respond to users' requests are called task execution costs. The average task execution costs of the proposed method, in comparison with the Tamilvizhi method and Power-Check, in these three considered scenarios, are shown in Fig. 8.

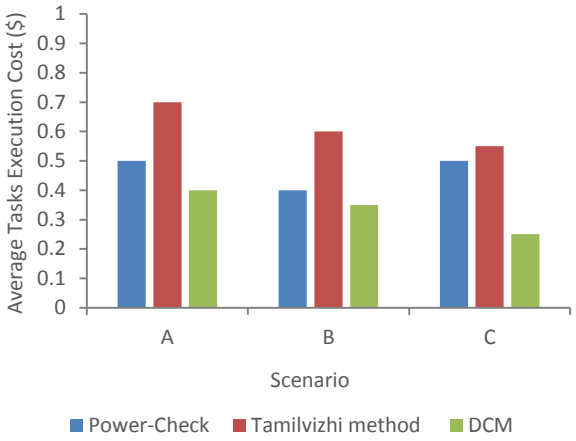


Fig. 8: Average Tasks Execution Costs

As shown in Fig. 8, the most proper host is selected for migration in the proposed method. Due to the exact prediction of failure, the possibility of failure and disturbance in tasks' performance is significantly reduced.

Also, the computational overhead is minimized, and the speed of the system is increased. Therefore, the method declines the average task execution costs compared with the Tamilvizhi method and Power-Check by 44.71% and 44.16%, respectively.

Fig. 9 illustrates the proposed method's average energy consumption compared to the Tamilvizhi method and Power-Check in these three considered scenarios.

The proposed method employs an exact load balancing method to minimize migration time (Fig. 9). In the method, the proper host for migration is attentively selected.

Therefore, on average, the DCM fault tolerance algorithm's energy consumption is optimized by 30% compared with that of other methods. Also, the energy consumption of the Tamilvizhi method is higher compared with that of others.

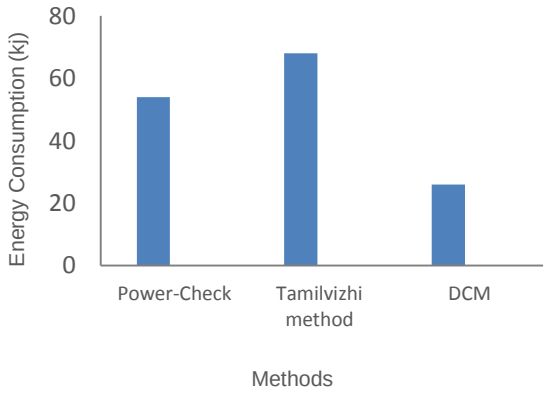


Fig. 9: Energy Consumption in HPC Systems.

Conclusion

In recent years, cloud computing has become a popular computing technology in all industries and provides more benefits than other technologies. One of the main challenges of cloud computing is fault tolerance, which avoids restarting the system and declines operational costs and energy consumption. In this paper the DCM method to enhance FTDaemon is proposed. In the proposed method, the Daemon Fault Tolerance technique has been enhanced, and COA-MMT has been utilized for load balancing.

The method consists of four modules, which are used to determine the host state. To predict and prevent failures, the proposed method utilizes these parameters: CPU temperature, CPU utilization, number of users' requests, voltage, and fan speed parameters. Based on evaluations and simulations, the proposed method is significantly optimized in task execution costs, job makespan time, and response time and declines the energy consumption compared with the Tamilvizhi method and Power-Check.

Author Contributions

H. Barati and A. Mehranzadeh conceptualized the research. H. Jahanpour designed the experiments and collected the data and she carried out the data analysis. H. barati and H. Jahanpour validated the results. H. Jahanpour wrote the manuscript. H. Barati and A. Mehranzadeh reviewed and edited the manuscript.

Acknowledgment

The authors would like to thank Dezful Branch, Islamic Azad University.

Conflict of Interest

The author declares that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancy have been completely observed by the authors.

Abbreviations

IT	Information Technology
HPC	High-Performance Computing
DCM	Daemon-COA-MMT
QoS	Quality of Service
SLA	Service Level Agreement
VM	Virtual Machine
SaaS	Software as a Service
IaaS	Infrastructure as a Service
HaaS	Hardware as a Service
PSO	Particle Swarm Optimization
Bee-MMT	Bee colony Algorithm-Minimal Migration Time
MMT	Minimal Migration Policy Time
CPU	Central Processing Unit
FTDaemon	Daemon's fault tolerance
COA-MMT	Cuckoo Optimization Algorithm-Minimum Migration Time
ELR	Egg Laying Radius
ADA	Adaptive Dragonfly algorithm

References

[1] M. Vaishnnave, K.S. Devi, P. Srinivasan, "A survey on cloud computing and hybrid cloud," Int. J. Appl. Eng. Res., 14: 429-434, 2019.

[2] M.U. Bokhari, Q. Makki, Y.K. Tamandani, "A survey on cloud computing," Big Data Analytics: 149-164, 2018.

[3] F.A. Ibrahim, E.E. Hemayed, "Trusted cloud computing architectures for infrastructure as a service: Survey and systematic literature review," Computers & Security, 82: 196-226, 2019.

[4] A.M. Caulfield, E.S. Chung, A. Putnam, H. Angepat, D. Firestone, J. Fowers, et al., "Configurable clouds," IEEE Micro, 37(3): 52-61, 2017.

[5] F. Zafar, A. Khan, S.U.R. Malik, M. Ahmed, A. Anjum, M.I. Khan, et al., "A survey of cloud computing data integrity schemes: Design challenges, taxonomy and future trends," Computers & Security: 65, 29-49, 2017.

[6] K. O'brien, I. Pietri, R. Reddy, A. Lastovetsky, R. Sakellariou, "A survey of power and energy predictive models in HPC systems and applications," ACM Computing Surveys (CSUR), 50(3):1-38, 2017.

[7] M.A. Netto, R.N. Calheiros, E.R. Rodrigues, R.L. Cunha, R. Buyya, "HPC cloud for scientific and business applications: taxonomy, vision, and research challenges," ACM Computing Surveys (CSUR), 51(1): 1-29, 2018.

[8] A. Pradhan, S.K. Bisoy, P.K. Mallick, "Load Balancing in Cloud Computing: Survey," Innovation in Electrical Power Engineering, Communication, and Computing Technology: 99-111, 2020.

[9] M.R. Mesbahi, A.M. Rahmani, M. Hosseinzadeh, "Reliability and high availability in cloud computing environments: a reference

roadmap," Human-centric Computing and Information Sciences, 8(1): 20, 2018.

[10] M.N. Cheraghlou, A. Khadem-Zadeh, M. Haghparast, "A survey of fault tolerance architecture in cloud computing,"Journal of Network and Computer Applications, 61: 81-92, 2016.

[11] A. Rezaeipanah, M. Mojarad, A. Fakhari, "Providing a new approach to increase fault tolerance in cloud computing using fuzzy logic," International Journal of Computers and Applications: 1-9, 2000.

[12] Q. Lin, K. Hsieh, Y. Dang, H. Zhang, K. Sui, Y. Xu, et al., "Predicting Node failure in cloud service systems. in Proc. the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering: 480-490, 2018.

[13] A.A. Shaikh, S. Ahmad, "Fault tolerance management for cloud environment: a critical review," International Journal of Advanced Research in Computer Science, 9(Special Issue 2): 34, 2018.

[14] A. Hota, S. Mohapatra, S. Mohanty, "Survey of different load balancing approach-based algorithms in cloud computing: a comprehensive review," Computational intelligence in data mining: 99-110, 2019.

[15] P. Kumar, R. Kumar, "Issues and challenges of load balancing techniques in cloud computing: A survey," ACM Computing Surveys (CSUR), 51(6): 1-35, 2019.

[16] M. Kumar, S.C. Sharma, "Dynamic load balancing algorithm to minimize the makespan time and utilize the resources effectively in cloud environment," International Journal of Computers and Applications, 42(1), 108-117, 2020.

[17] K. Pan, J. Chen, "Load balancing in cloud computing environment based on an improved particle swarm optimization," in Proc. 2015 6th IEEE International Conference on Software Engineering and Service Science (ICSESS): 595-598, 2015.

[18] F. Abazari, M. Analoui, H. Takabi, S. Fu, "MOWS: multi-objective workflow scheduling in cloud computing based on heuristic algorithm," Simulation Modelling Practice and Theory, 93: 119-132, 2019.

[19] M. Abd Elaziz, S. Xiong, K.P.N. Jayasena, L. Li, "Task scheduling in cloud computing based on hybrid moth search algorithm and differential evolution," Knowledge-Based Systems, 169: 39-52, 2019.

[20] Y.L. Huang, Z.X. Li, "A GA-based resource management algorithm for smart living applications requiring intensive computing power," in Proc. 2017 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW): 259-260, 2017.

[21] S.S. Abdhullah, K. Jyoti, S. Sharma, U.S. Pandey, "Review of recent load balancing techniques in cloud computing and BAT algorithm variants," in Proc. 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom): 2428-2431, 2016.

[22] S.M. Ghafari, M. Fazeli, A. Patooghy, L. Rikhtechi, "Bee-MMT: A load balancing method for power consumption management in cloud computing," in Proc. 2013 Sixth International Conference on Contemporary Computing (IC3): 76-80, 2013.

[23] I.P. Egwutuoha, S. Chen, D. Levy, B. Selic, R. Calvo, "Energy efficient fault tolerance for high performance computing (HPC) in the cloud," in Proc. 2013 IEEE Sixth International Conference on Cloud Computing (CLOUD): 762-769, 2013.

[24] I.P. Egwutuoha, S. Chen, D. Levy, B. Selic, R. Calvo, "A proactive fault tolerance approach to High Performance Computing (HPC) in the cloud," in Proc. 2012 Second International Conference on Cloud and Green Computing (CGC): 268-273, 2012.

[25] R.R. Chandrasekar, A. Venkatesh, K. Hamidouche, D.K. Panda, "Power-check: An energy-efficient check pointing framework for HPC clusters," in Proc. 2015 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid): 261-270, 2015.

[26] M. Yakhchi, S.M. Ghafari, S. Yakhchi, M. Fazeli, A. Patooghi, "Proposing a load balancing method based on Cuckoo Optimization Algorithm for energy management in cloud computing infrastructures," in Proc. 2015 6th International Conference on Modeling, Simulation, and Applied Optimization (ICMSAO): 1-5, 2015.

[27] T. Tamilvzhi, B. Parvathavarthini, "A novel method for adapive fault tolerance during load balancing in cloud computing," Cluster Computing, 22(5): 10425-10438, 2019.

[28] P. Neelima, A.R.M. Reddy, "An efficient load balancing system using adaptive dragonfly algorithm in cloud computing," Cluster Computing, 23: 2891.2899, 2020.

[29] T.D. Devi, A. Subramani, P. Anitha, "Modified adaptive neuro fuzzy inference system based load balancing for virtual machine with security in cloud computing environment," Journal of Ambient Intelligence and Humanized Computing, 1-8, 2020.

[30] L. Kong, J.P.B. Mapetu, Z. Chen, "Heuristic load balancing based zero imbalance mechanism in cloud computing," Journal of Grid Computing, 18(1): 123-148, 2020.

Biographies

Hoda Jahanpour received her B.Sc. degree in Computer Engineering from Dezful Branch, Islamic Azad University, Dezful, Iran, in 2014. Furthermore, she received her M.Sc. degree in computer systems Architecture from Dezful Branch, Islamic Azad University, Dezful, Iran, in 2016. Her major research interests include Distributed Computing and cloud computing.

Hamid Barati is an Assistant Professor in the Department of Computer Engineering at Dezful Branch, Islamic Azad University, Dezful, Iran. He received his B.S. degree in Computer Hardware Engineering, M.S. degree in Computer Systems Architecture Engineering and Ph.D. degree in Computer Systems Architecture Engineering in 2005, 2007 and 2015 respectively. Currently he is Faculty of Islamic Azad University, Dezful Branch, Iran. His major research experiences and interests include mobile ad hoc networks, interconnection networks and energy-efficient routing and security issues in wireless sensor networks.

Amin Mehranzadeh received the M.Sc. and Ph.D. degrees in Computer Engineering. He is currently an Assistant Professor in Computer Engineering Department at Azad University of Dezful. His research interest is Distributed Computing, Cloud Computing, Embedded Systems and Network-on-Chip systems including Performance and Cost Improvement in Routing and Arbitration of various types of NoC systems. Recently, he has started carrying out research in Deep Neural Network with his team which consists of M.Sc. and Ph.D. students.

How to cite this paper:
H. Jahanpour, H. Barati, A. Mehranzadeh, "An Energy Efficient Fault Tolerance Technique Based on Load Balancing Algorithm for High-Performance Computing in Cloud Computing," Journal of Electrical and Computer Engineering Innovations, 8(2): 169-182, 2020.

DOI: 10.22061/JECEI.2020.7219.371

URL: http://jecei.sru.ac.ir/article_1467.html

Journal of Electrical and Computer Engineering Innovations
(JECEI)
Journal homepage: http://www.jecei.sru.ac.ir

Research paper

NSE-PSO: Toward an Effective Model Using Optimization Algorithm and Sampling Methods for Text Classification

R. Asgarnezhad[1], S.A. Monadjemi[2,*], M. Soltanaghaei[1]

[1]*Department of Computer Engineering, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran*

[2]Faculty of Computer Engineering, University of Isfahan, Isfahan, Iran, and Senior Lecturer, School of continuing and lifelong education, National University of Singapore, Singapore, 119077.

Article Info

Article History:
Received 7 September 2019
Reviewed 14 December 2019
Revised 25 February 2020
Accepted 27 May 2020

Keywords:
Text classification
Sampling technique
Feature selection
Optimization algorithm
Twitter

*Corresponding Author's Email Address:
monadjemi@eng.ui.ac.ir

Abstract

Background and Objectives: With the extensive web applications, review sentiment classification has attracted increasing interest among text mining works. Traditional approaches did not indicate multiple relationships connecting words while emphasizing the preprocessing phase and data reduction techniques, making a huge performance difference in classification.

Methods: This study suggests a model as an efficient model for sentiment classification combining preprocessing techniques, sampling methods, feature selection methods, and ensemble supervised classification to increase the classification performance. In the feature selection phase of the proposed model, we applied n-grams, which is a computational method, to optimize the feature selection procedure by extracting features based on the relationships of the words. Then, the best-selected feature through the particle swarm optimization algorithm to optimize the feature selection procedure by iteratively trying to improve feature selection.

Results: In the experimental study, a comprehensive range of comparative experiments conducted to assess the effectiveness of the proposed model using the best in the literature on Twitter datasets. The highest performance of the proposed model obtains 97.33, 92.61, 97.16, and 96.23% in terms of precision, accuracy, recall, and f-measure, respectively.

Conclusion: The proposed model classifies the sentiment of tweets and online reviews through ensemble methods. Besides, two sampling techniques had applied in the preprocessing phase. The results confirmed the superiority of the proposed model over state-of-the-art systems.

Introduction

Regarding the explosion of information on the Internet, it is difficult to make decisions based on reviews, tweets, etc. People purchase products on the Internet and give their reviews about them every second. These reviews affect the financial statements in companies noticeably [1] [2] [3] [4] [5] [6] [7] [8] [9]. The main problem is that reviews are in natural language. There is a big gap between reviews in natural language and applications that use structured data. Sentiment Classification is a key tool in this field.

The Sentiment Classification problem is an attractive field in text mining. This field extracts the reviews from the unstructured data on the Internet to organize reviews into two or three classes [10] [11]. Twitter-Sanders-Apple (TSA) generated by Sanders Analytics. Identifying challenges can be considered in Twitter Sentiment Classification consist of classification

accuracy, data sparsity, neutral tweets, and linguistic representational; also, tweets are concise. These problems increase the review classification error. The reason behind that problem is the lack of beneficial relationships between words and sampling techniques applied. Twitter Sentiment Classification is different from other domains in Sentiment Analysis. Almost all movie reviews tend predominantly to be positive or negative; also, no sentiment likely belongs to a feature in a sentence of reviews. With the growing amount of reviews, the effect of review quality on the preprocessing techniques becomes undeniable.

Preprocessing has a principal role in Sentiment Classification. It showed that traditional approaches could not provide enough information for Natural Language Processing analyses. Traditional approaches will add unnecessary complexity; in contrast, words are well indicators for sentiment polarity detection. Traditional BOW did not record multiple relationships between words; hence, we add n-grams to the Bag of Word approach to extract features based on the relationships of the words. Many works combined Machine Learning (ML) algorithms with n-grams. The significant results which were more optimal rather than base classifiers achieved [12] [13]. Specifically, we use n-gram features and sampling techniques in preprocessing steps.

Sentiment Classification approaches can be divided into Lexicon-based, ML [14] [15], and Hybrid approaches. ML aims to optimize an algorithm to increase the system performance using examples and experiences in the past. The ML approaches exist based on three methods like supervised, unsupervised, and semi-supervised. The unavailable labeled dataset is a significant drawback for the supervised methods because they obtain the words with a certain domain. Two critical stages in this context are feature and classifier selection for determining the performance of classification. It has revealed that the ML algorithms like Naive Bayes (NB), support vector machine (SVM) [16], and maximum entropy utilized successfully in many types of research. The current author applied the supervised approaches in conjunction with ensemble methods as different alternatives herein. The acquisition of the domain for words relating to a domain corpus is the main benefit of the lexicon-based approach. A hybrid model combines the services of both them to improve the performance of classification. It is obtaining robust accuracy, and endurance of the two mentioned approaches.

The supervised methods are simple and ordinary; in contrast, ensemble methods like bagging obtain more accurate results. We apply boosting, stacking, and voting as other alternatives in this study. A bagging method assigns equal weights to embedded classifiers, but a boosting method gives a particular importance to each embedded classifier. Their results were good enough; therefore, we add sampling techniques and n-grams to our model to improve more.

In the current study, the effect of different combinations of preprocessing, sampling, Particle Swarm Optimization (PSO), and ensemble techniques investigated on the performance of classification. It is distinct from the existing studies due to employing these different combinations; also, both binary and multi-class classifications are applied. After applying a series of operations in preprocessing phases, we form n-gram features with two sampling techniques to improve the performance of both binary and multi-class classifications. Two weighting mechanisms, term frequency-inverse document frequency (IDF) and term frequency (TF) employed to form the word vector. Two supervised methods, and Ensemble classification method employed. To evaluate the proposed model, TSA datasets considered. However, the Twitter dataset is not available, except Sanders. It seems that weighting feature mechanisms obtained different results on the datasets. As shown in our study, the highest precision obtained through TF mechanism on the TSA2 datasets; whereas, the highest precision was obtained through the TFIDF mechanism on the TSA3 dataset. It also appears that bootstrapping sampling, PSO algorithm achieved higher results on the three datasets. The highest results achieved through our optimized model for the boosting method on the datasets. Experiments showed that our independent-domain approach can improve the classification performance and outperform the existing traditional techniques [4] [17] [18] [19].

We used a sampling technique to emphasize that our model is different and significant in binary and multi-class classifications. Also, concerning more sophisticated methods, this model is simple. Here, our contributions to this research shortened as follows:

- Inspiring Data reduction through sampling technique
- Choosing the best feature by the PSO algorithm
- Improving the performance of the classification model relating ensemble methods

The excess of this article organized as follows: Related work contains a summary of the works. Next, the proposed model presented and evaluated through the experiments explained. Conclusively, the article ended in this article

Related work

Sentiment Classification is attracting considerable attention due to its applications in the new year. Several works proposed to improve the classification performance on the known datasets. Those works differ from each other in the preprocessing, classifiers, and applied datasets. Here, we explain some of these works

from 2014 to 2020 on Twitter.

In 2014, Da Silva et al. posed a unique approach for many applications in Sentiment Analysis [3]. It seems that the supervised ML methods obtained high accuracy among other methods; hence, we use the supervised methods. Note that supervised methods have a distinct drawback, an available labeled data. These methods have more computational complexity compared to unsupervised methods. Ensemble Classification methods are other alternatives.

After one year, Tripathi and Naganna in [13] attempted to make a different preprocessing scheme for investigating the behavior of NB and SVM classifiers without sampling technique. They found that n-grams obtained higher results for sentiment analysis and the best accuracy achieved by bigrams.

Tripathy et al. used the combination of TFIDF to produce a digit matrix from the text in 2016 [20]. After one year, Pandey et al. proposed a novel clustering method using a cuckoo algorithm on the TSA dataset in 2017 [4].

In 2018, Trupthi et al. [19] investigated the effective topic modeling methodology Latent Dirichlet Allocation to extract the keywords in a clustering manner. Next, they applied the keywords using the fuzzy c-means approach on the twitter dataset. Vashishtha and Susan in [21] proposed a system to classify the posts in social media through fuzzy rules. They offered a new system, which combines nine fuzzy rules with techniques for Word Disambiguation. They reached 58.9, 59.7, and 68.6% of recall, precision, and f1-score values on the TSA3 datasets, respectively. After one year, Tripathi et al. [23] suggested a novel Map-Reduce based K-means to cluster the large scale data.

The present researchers in 2018 [22] proposed a model named SFT for Twitter Sentiment Classification in 2018. The goal of our model was to investigate the role of weighting feature techniques in Sentiment Classification using supervised methods on the Twitter data set. The applied classifier in the current article based on the SFT model in our previous article. The previous descriptions revealed that no work combined the sampling technique with the ensemble method. Therefore, the current study proposed a novel Sentiment Classification model. In 2020, Abbas et al. [24] offered a classification model with four classifiers, and varying techniques to form a single ensemble classifier. They gained an accuracy of 82.2% on Twitter. Also, Jiang et al. [25] develop a novel Neural Network-based model to conduct the aspect-level Sentiment Classification tasks. Naseem et al. [26] shown a transformer-based method for Sentiment Analysis and applied deep learning and the bidirectional Long Short Term Memory network through omitting noise to heighten the classification performance. They reached an accuracy of 96.2% on airline datasets.

Samad et al. [27] investigated the effect of seven scenarios for text processing on Twitter. Their experiments revealed adverse effects on Sentiment Classification of two common text processing steps: 1) stop word removal; 2) averaging word vectors to represent individual tweets. Word selection from context-driven word embedding showed that only the ten most essential words in Tweets cumulatively produce over 98% of the maximum accuracy.

Sharma and Jain in [28] presented the usage of various ML techniques for collecting tweets and assessing sentiments. After gathering data from twitter, they applied preprocessing and feature extraction stages for the text data. Selection methods based on correlation used and ML classifiers to confirm which classifier gives better results. They obtained an accuracy of 88.2% on the Cambridge Analytica dataset.

The current authors in 2020 [18] proposed a new model based on fuzzy analytic hierarchy on Twitter, namely FAHPBEP. The highest f-measure obtained 90.88 and 90.01% for TSA2 and TSA3, respectively. Also, they in [17] suggested a hybrid model based on ensemble methods on Twitter, namely NSET. The highest f-measure obtained 93.52 and 89.64% for TSA2 and TSA3, respectively. We found that using preprocessing techniques in conjunction with ensemble classification methods may enhance the performance results. We believe that these are unseen combinations of ensemble classification methods. Applying proper alternatives that have not considered in the literature, our obtained results may lead to more accurate performance.

The proposed Model

This article introduces a hybrid model to document-level Sentiment Classification. This approach explores multi-class classifications. In multi-class types, sentiments classified into three classes or more. The model investigates the effects of sampling technique, n-grams, and PSO algorithm using ensemble methods for Sentiment Classification. It seems that the combination of sampling methods, weighing schema, and PSO algorithm can improve the classification performance. The proposed model exploits supervised methods to handle the document-level multi-class Sentiment Classification. A set of operations considered for the preprocessing phase.

After studying many classification algorithms [30], we applied the SVM as a baseline classifier. Ensemble classifiers and parameter optimization were two important of our model which used herein. The results indicate that the proposed model outperforms others' works in text classification. Also, the model gives a higher performance through sampling and PSO

algorithms. Our proposed model, namely NSE-PSO, investigates the effects of n-grams, weighting feature mechanisms, sampling techniques, and PSO algorithm using ensemble classification methods, in a stepwise manner. Fig. 1 exposes the phases of our proposed NSE-PSO model.

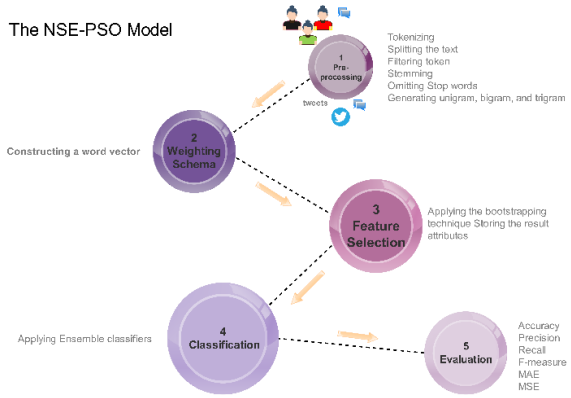


Fig. 1: The phases of the proposed model.

First, the unimportant and worthless characters tokenized and then removed from the text. Next, we employ n-gram features and two weighting feature mechanisms for forming the word vector. It produces to the dataset reduced effectively. Attribute subset selection techniques used to improve classification performance. The PSO algorithm is applied to select the best features. Finally, the classification phase performed by supervised and ensemble classification methods. An extensive range of comparative experiments done on the datasets to investigate the effectiveness of our proposed model. By experiments, the state-of-the-art result obtained. We used the linear SVM as a base classifier in bagging and boosting methods. The comparison among the best results in the literature and our obtained results handled. We employed a type of sampling technique to emphasize that the obtained results are different and significant. In contrast to other sophisticated methods, our hybrid model outperforms good enough in this context.

Here, the pseudo-code of our model expressed as:

Pseudo-code for the NSE-PSO model

Pseudo-code for the preprocessing
1: **Input:** A dataset;
2: **Output:** A classification model for a set of words;
3: **For each** document **in** dataset **do**
4: Tokenizing characters, words, and useless tokens
5: Splitting the text into a sequence of tokens
6: Filtering tokens based on their length
7: Stemming word via Porter Algorithm
8: Omitting Stop words based on the stop word list
9: Generating unigram, bigram, and trigram
10: **End.**

Pseudo-code for the weighting schema
1: **Input:** A set of words;
2: **Output:** A set of word vectors;
3: **For each** set of words **in** Input **do**
4: Constructing a word vector based on TF & IDF schemas
5: **End.**

Pseudo-code for the sampling techniques
1: **Input:** A set of word vectors;
2: **Output:** A set of word vectors;
3: **For each** word vector **in** Input **do**
4: Applying the sampling techniques and storing the result attributes.
5: **End.**

Pseudo-code for the Particle Swarm Optimization and Classification
1: **Input:** Training set;
2: **Output:** A composite model and confusion matrix;
3: Maximum iteration, Population size, inaction weight;
4: Generate initial population;
5: **For each** step **in** maximum iteration **do**
6: Perform the ensemble model for each set of parameters (Classifying and calculating the fitness function);
7: Update the fitness function;
8: Stopping and obtaining the optimal parameters;

In the following, the phases of the NSE-PSO model are discussed in detail.

A. Preprocessing and Weighting Mechanism Phases

The preprocessing stage consists of Tokenization, Filter-Token, Stemming, Filtering Stop Word, and N-grams. Two weighting feature mechanisms used to create word vectors. The TF defined as total occurrences of the word *t* in document *d* divided by a total volume of the comments happening in document *d*. The TFIDF defined by Manning et al. in (1) as,

$$TFIDF = (TF) \times Log(N / F_t) \quad (1)$$

where, TF is the frequency of word t in document *d*, N is the volume of documents in a group, and Ft is the volume of documents in a collection containing word *t* [29].

B. Sampling and Feature Selection Phases

Here, we describe the sampling techniques and PSO algorithm in our model. These techniques caused to select the best features through data reduction. Stratified and bootstrapping samplings are two of the sampling techniques, which used to attain better performance. In stratified sampling, the folds of the training set stratified. The class distribution for tuples in a fold is similar to the initial data. It enables the algorithm to preserve the distribution of the training set. Bootstrapping sampling creates a bootstrapped sample from the dataset. This type of sampling may not have all unique examples; hence it is different from other sampling techniques. We use both samplings in our model, but bootstrapping sampling with replacement achieves higher performance. We apply to bootstrap sampling in all of the experiments. The obtained results show that bootstrapping sampling gives a higher performance on the text input.

To produce an experimental distribution of sampling, a sampling structure of scores was applied. Sampling with or without replacement is the first portion. The size

of the sample is the second portion. When a full sampling with replacement happens, the bootstrap is defined as a procedure herein. This bootstrap applies the dimensions of the sample equal to the size of the original collection. In the second portion, a taxonomy determined by extending the sizes of the possible sample. Here, a distribution of the sampling with dimensions more extensive than the dimension of the original data not be applied. Nevertheless, this event did not consolidate for the samples without replacement. Therefore, this proposal only admits expansion for the bootstrap. The tuple likely selects and adds to the training set again when we choose each tuple of the input set. In each step of iterations, all examples have an equal probability of being selected. When an example chooses, it remains a candidate for further selection and determines again in the next step. It is undeniable that a sample with replacement can have the same examples. Therefore, it used to create a sample that is greater in size than the original one. It appears that bootstrapping performs better than the former one. The idea behind bootstrapping is using the data as an input set for approximating the sampling distribution. It creates an enormous volume of samples named bootstrap samples. The sample outline calculated for each sample of bootstrap [30] [31]. Here, some notations for utility described. Assume a parameter for a population θ is as a target herein. An random sample with size *n* produces the data $(X_1,X_2,...,X_n)$. Assume θ is a sample created from the dataset. The distribution of θ with large n can be bell-shaped with a center θ and standard deviation $\sigma/\sqrt{n}$ for each sample, where the positive volume depends on two factors like the population, and the type of statistic θ. There exist technical complexity for standard deviation, when θ is median or correlation of sample. Therefore, bootstrap assigns a bypass. Assume θ_B is considered as a quantity for presenting the same statistic, which produces on a bootstrap sample of $(X_1,X_2,...,X_n)$.

With the limitation of $(n\to\infty)$, the distributions of θ_B were bell-shaped with θ as the center and the corresponding standard deviation $\sigma/\sqrt{n}$. Thus, the distribution $\theta_B-\theta$ will be the distribution $\theta-\theta$. This distribution is named the bootstrap Central Limit Theorem (CLT) [32] [33].

It also observed that with a limiting distribution of the sampling for a mathematical function that does not include population unknowns, bootstrap distribution assigns a better conjecture than the CLT. If the procedure is $(\hat{\theta}_B-\hat{\theta})/SE$, where SE considered as a sample estimation of the standard error of $\hat{\theta}$, the limiting sampling distribution will be standard normal.

Here, $\theta=\mu$ is the population means, $\theta=\overline{X}$ is the sample mean, σ is population standard deviation, and s is sample standard deviation considered, which produced from the original dataset. Also, s_B is the sample standard deviation, which calculated on a bootstrap sample. Next, the sampling distribution of $(\overline{X}-\mu)/SE$, with $SE=\sigma/\sqrt{n}$, will be estimated through the bootstrap distribution of $(\overline{X_B}-\overline{X})/SE$, where $\overline{X_B}$ is bootstrap sample means, and $SE=s/\sqrt{n}$. Likewise, the sampling distribution of $(\overline{X}-\sigma)/SE$, where $SE=s/\sqrt{n}$, will be assessed through the bootstrap distribution of $(\bar{X}_B-\bar{X})/SE_B$, where $SE=s/\sqrt{n}$. Here, the description of the approximating standard error of sample evaluation for utility is of concern. We assume that the information investigated regarding the population parameter of θ, where $\bar{\theta}$ is a sample estimator of θ based on a stochastic sample has size *n*. To estimate the standard error for $\hat{\theta}$, a bootstrap approach is of concern: calculate $(\theta_1^*,\theta_2^*,...,\theta_N^*)$, through the equivalent relation for $\hat{\theta}$, exactly with N numbers of different bootstrap samples. A primary recommendation for the size N could be $N=n^2$, unless n^2 be too large. In that case, it could be reduced to an acceptable size, say $n\log_e^n$. So, $SE_B(\hat{\theta})$ defined as (2),

$$SE_B(\hat{\theta})=\sqrt{\frac{\sum_{i=1}^{N}(\theta_i^*-\hat{\theta})^2}{N}} \quad (2)$$

It revealed that more instances could exploit more useful information about the dataset. Therefore, it may consider a novel example in the dataset, which is noticeable for classification. That is why bootstrapping sampling has become a useful tool in our model.

PSO is a global technique for optimization problems, proposed by Kennedy et al. [34]. A group of particles represented by x_{ij} search of the solution space to obtain the best solution. Each particle has a place, velocity, and memory to preserve its best position. The rate represented by v_i. This technique is well-known for ease of implementation, convergence speed, and few parameters to adjust, whereas a particle may converge on a suboptimal solution. Each candidate's answer could interpret as a particle with a place in the state space. With particle movement in that space, the optimal solutions emerged. Within a change, each particle refreshes its location and speed according to its neighbors. The best previous place of the particle registered as the individual best p-best, and the best location captured through the population of g-best. Cognitive and social scaling parameters are known as C_1, and C_2. So, the obtained optimal answers through

renewing the speed and place of each particle are of concern, (3):

$$V_{ij}^{r+1} = V_{ij}^{r+1} + C_1 rand_1 (p - best_{ij} - X_{ij}^{r}) + C_2 rand_2 (g - best_{ij} - X_{ij}^{r})$$
$$X_{ij}^{r+1} = X_{ij}^{r} + V_{ij}^{r+1} \quad (3)$$

Ultimately, the algorithm stopped by a predefined criterion like a proper fitness amount or a maximum number of iterations. Here, the PSO chooses particles randomly to explore the optimal particle. Per particle is represented as an m-dimensional point/node. The AdaBoost classifier is applied to evaluate effectiveness by the cross-validation technique. PSO investigates the determination of possible subsets to achieve the most significant accuracy. When the efficiency of AdaBoost converges, the iterations end. As mentioned before, PSO configured with population size, inaction weight, and maximum iteration to generate the initial population. The most useful alternative for the local and global parameters examined by evaluating the fitness function of each particle. Then, the speed and location of each particle updates for the values of the fitness converges. Finally, the global best applied for the training of the AdaBoost.

C. Classification phase

Ensemble classification methods applied to obtain better performance. An ensemble combined weak learners to produce a strong learner. It mainly requires more computation to evaluate the prediction; whereas, a single model requires fewer calculations. However, ensemble techniques are a tool for improving the poor performance of base learning algorithms. When the ensemble method trained, it determines a hypothesis. This hypothesis did not contain only the models used for its training; hence, ensembles are more flexible in their functions. However, it can end in over-fitting over them. These methods often tend to yield better results when used in diverse models. Almost all ensemble methods use a diversity of the models to improve the poor performance for each single learners. Among the mentioned classifiers, AdaBoost outperforms the others better. It found that ensemble methods, especially AdaBoost, employed to increase the precision and accuracy through combining a series of individual classifiers. SVM is an ML technique based on the statistical learning concept, which performed well in text classification applications. In the current study, we use a boosting model in conjunction with SVM as base learners. We focus on using sampling and ensemble methods as a base learner for classification. Here, several methods incorporated in the proposed model. In comparison with the literature, our model achieves higher results. That is because the usage of ensemble methods in conjunction with sampling techniques and PSO algorithms gives the best features.

D. Data sets and Evaluation Phase

There are a few available and free resources on Twitter. None of the existing datasets on Twitter are free, except the Sanders dataset. Hence, two datasets of TSA applied for training and testing experiments. The used datasets were generated by Sanders Analytics. Two datasets of the TSA are accessible at http://www.sananalytics.com.

Here, measures for evaluating Sentiment Classification had introduced. P and N are the numbers of positive and negative tuples. TP refers to the positive tuples that have been labeled by the classifier correctly. TN refers to the number of true negatives. FP is the negative tuples that have labeled incorrectly as positive. FN is the positive tuples that have mislabeled as negative. Accuracy is the sum of actual tuples that classified TP and the number of TN relative to the total number of classified instances. The precision state as the percentage of tuples that have labeled as positive and actual. Recall refers to the percentage of tuples that are labeled positive. F-measure combines precision and recall into a single measure [12] [35] [29]. F-measure comes from a weighted harmonic mean of precision and recall. Also, mean absolute error (MAE) and root absolute error (RAE) for error evaluation employed. These measures computed in (4) to (9).

$$Accuracy = \frac{TP + TN}{P + N} \quad (4)$$

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{P} \quad (6)$$

$$F - measure = \frac{2PR}{P + R} \quad (7)$$

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} \quad (8)$$

$$RAE = \frac{\sum_{i=1}^{n} |x_i - \overline{x_i}|}{\sum_{i=1}^{n} |y_i - \overline{y_i}|} \quad (9)$$

Results and Discussion

In all experiments, we assume epsilon and the kernel type of the SVM is one and C-SVC, respectively. The remaining parameters for SVM optimized in all experiments. Also, the evaluation model of the kernel NB supposed greedy, with the number of kernels equals 10. During the experiments, the 10-fold cross-validation utilized again, and performance evaluation parameters calculated. In some cases, the random seed used to have

different examples. By changing the value of this parameter, we can change the way examples are randomized. Here, a maximum number of iterations for the PSO algorithm considered 100. Implementation of R programming used to conduct the experiments.

Experiment I. Here, we investigate the effect of the PSO algorithm for choosing the best features on the TSA2 dataset. Table 1 shows the obtained results for our model. The highest accuracy, f-measure, and precision values obtained 92.61, 94.49, and 97.33%, respectively. It also seems that the performance of used model using IDF mechanism is higher than that of the TF on the TSA2 dataset. The reason is that each example in this dataset has one or two sentences. It found that the highest performance for the proposed model obtained relating to PSO, IDF, sampling, and bigrams in general. It observed that bootstrapping sampling can improve the performance measure in most cases. This is due to bootstrapping sampling produces a dataset greater than the original dataset. Here, the accuracy and f-measure improved rather than without sampling approximately 10 and 12%, respectively. Fig. 2 presents the highest performance of the experiment I.

Table 1: The obtained results of the proposed method in experiment I

Pre.		P	R	F	A	MAE	RAE
The results after using Sampling and PSO							
IDF	**N=1**	96.50	91.26	90.88	92.19	0.1334	0.2354
	N=2	97.33	91.82	94.49	92.61	0.1365	0.2357
	N=3	96.77	90.48	93.52	91.77	0.1365	0.2312
TF	**N=1**	90.82	86.13	88.41	87.82	0.1312	0.2343
	N=2	95.91	90.81	93.29	91.99	0.1342	0.2312
	N=3	95.49	88.81	92.02	91.37	0.1312	0.2343
No-sample		86.51	79.27	82.73	82.80	0.1256	0.2358

Note: Pre= Preprocessing, P=Precision, R=Recall, F=F-measure, A=Accuracy

Experiment II. We try to investigate the effect of the NSE-PSO on the TSA3 dataset. The AdaBoost method with SVM applied. Table 2 presents the obtained results of this experiment. It revealed that the highest accuracy, precision, recall, and f-measure rates obtained using IDF and PSO. The achieved f-measure of our model were more excellent than 96%; whereas, without sampling were 86.60%. The best-obtained accuracy was 88.75% through the IDF mechanism, unigrams, PSO, and sampling. The highest precision relating PSO, sampling, and trigrams was 95.33%; whereas, the f-measure using sampling, PSO, and bigrams gets through 96.23%. All best performances of the NSE-PSO without PSO achieved applying the TFIDF, bigrams, and unigrams. It appears that the usage of n-grams and PSO increases the results of the usage of sampling.

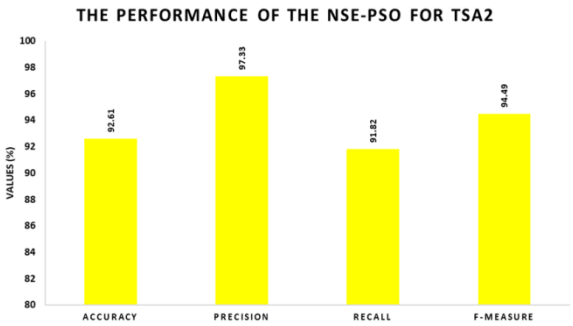


Fig. 2: The comparison between the obtained results using PSO algorithm on the TSA2 dataset.

Table 2: The obtained results of the proposed method in experiment II

Pre.		P	R	F	A	MAE
The results after using Sampling and PSO						
IDF	**N=1**	95.33	97.16	96.23	88.75	0.1243	0.2225
	N=2	93.76	96.12	94.92	87.74	0.1263	0.2234
	N=3	94.52	94.39	94.45	87.43	0.1254	0.2265
TF	**N=1**	89.51	88.03	88.76	82.58	0.1265	0.2276
	N=2	93.13	87.47	90.21	84.90	0.1222	0.2243
	N=3	93.41	86.91	90.04	84.50	0.1254	0.2243
No-sample		64.66	73.05	68.60	72.90	0.1256	0.2377

Note: Pre= Preprocessing, P=Precision, R=Recall, F=F-measure, A=Accuracy

It also shows that our model on the TSA3 dataset gives higher results using the IDF mechanism. The best general performance achieved by our optimized implementation for the boosting method on the TSA3. The highest f-measure achieved is 96.23%, which belongs to unigrams and PSO. The use of ensembles leads to additional computational costs, but the obtained accuracy is usually worthwhile. The improvement of accuracy is 15% in this experiment rather than the obtained results without sampling. The unnecessary features tokenized and filtered to apply the BOW technique. However, PCA applied as a data preprocessing, but the performance of classification does not improve. Fig. 3 exposes the best results of the NSE-PSO model before and after using sampling.

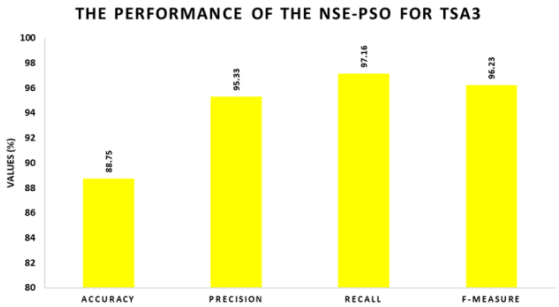


Fig. 3: The comparison between the obtained results using PSO algorithm on the TSA3 dataset.

Experiment III. Here, the best-obtained results from our proposed model and the best results in the literature on the datasets compared. Figs. 4 and 5 show the comparison among the NSE-PSO and three best works in this context.

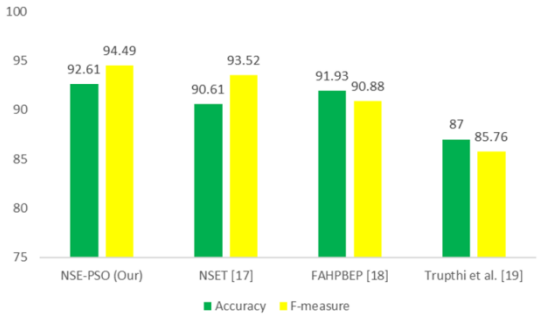


Fig. 4: The comprehensive comparison among our performance of experiment I on the TSA2 dataset.

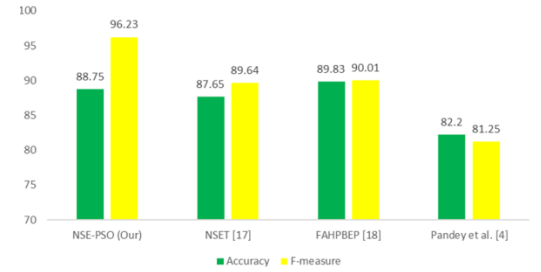


Fig. 5: The comprehensive comparison among our performance of experiment II on the TSA3 dataset.

It appears that bigram features can improve performance. Also, sampling can aim to obtain higher results. The experimental results emphasize that AdaBoost using sampling is substantially more accurate than other techniques in our tests. These results achieved using both TF and IDF mechanisms. It can also found that bigrams can improve the accuracy and f-measure for the two datasets. It seems that we decrease approximately the run time of ensemble methods through sampling. The striking comparison confirmed the advantage of our model over the state-of-the-art systems. It reveals that the proper use of sampling techniques and ensemble classification methods can improve the performance. It also seems that our model with using sampling, bigrams give higher results rather than other methods in the literature on the TSA datasets. According to Fig. 4, the highest accuracy of the proposed model obtained 92.61% on the TSA2 dataset; whereas, Trupthi et al. [19], the NSET in [17], and the FAHPBEP in [18] obtained 87, 90.61, 91.93%, respectively. This is due to that the ensemble method suggested in the NSET used only Adaboost in conjunction with SVM; whereas, we applied sampling and PSO besides boosting to select relevant word in the preprocessing phase. Also, the FAHPBEP and Trupthi et al. used fuzzy approach; whereas, we apply optimization algorithm. The highest f-measure of the proposed model obtained 94.49% on the TSA2 dataset; whereas, Trupthi et al. [19], the NSET in [17], and the FAHPBEP in [18] obtained 85.76, 93.52, 90.88%, respectively. Our obtained f-measure rate was approximately 1% greater than the best in the literature, i.e., 93.52%.

According to Fig. 5, we obtained the accuracy rate using the boosting method of 88.75% on the TSA3 dataset. Whereas Pandey et al. [4] obtained 82.20% of accuracy. They derived features from the k-means and cuckoo search, but we used a PSO algorithm. Despite using the bootstrap model and several classifiers, their framework was not more effective than our approach. It appears that IDF is the best weighting schema for TSA3. The reason for it is the few volume sentences in each example of the dataset. Also, the highest f-measure of our proposed model obtained 96.23% on the TSA3 dataset; whereas, Pandey et al. [4] obtained 81.25%. Notwithstanding applying the k-means model and cuckoo search, their framework was good enough and not more effective. Also importantly and opposition to the other works, higher classification f-measure achieved by our model. The accuracy rate reported of the NSET model in [17] on the TSA3 dataset was 87.65%; whereas, our implementation obtained an accuracy of 88.75%. The improvement of the f-measure rate in our model was approximately 6% greater than the best f-measure rate in the literature. The highest obtained f-measure of our model is 96.23% for the TSA3 dataset; whereas, the accuracy rate of the FAHPBEP in [18] gets through 90.01%. However, our accuracy rate obtained 88.75%, which was approximately 1% lower than the FAHPBEP. It revealed that the f-measure rate can be more significant rate for comparison on the TSA dataset. It revealed that bootstrapping gives the best results on all datasets. The benefit of utilizing ensemble methods is enhancing the classification performance; contrary to time, which it needs to cease the training phase of these methods is a disadvantage. Our concern was constructing an approach that has higher performance compared to the other works on three datasets. No work has been published on Sentiment Classification using the combination of sampling, n-grams, PSO algorithm, and ensemble methods. The proposed model not only applies the ensemble method and sampling but also utilizes a meta-classifier. It seems that sampling and PSO yield better results compared to these existing methods. The obtained results emphasize that AdaBoost using PSO algorithm is substantially more accurate than other applied techniques in our model. Improvement of f-measure calculated to estimate the performance of the present model on the datasets. The improvement measures are of concern in (10):

$$improvement\ of\ f-measure = $$
$$=(f\text{-}measure_{NSE-PSO} - f\text{-}measure)/f\text{-}measure_{NSE-PSO} \quad (10)$$

For the TSA2 dataset, the improvement of f-measure is 0.97%. For the TSA3 dataset, the progress of f-measure is 6.22%. The benefit of utilizing PSO algorithm is enhancing the classification performance; in contrast, time that needs to cease the training phase of the method is a disadvantage. We used sampling techniques to decrease this time. Our goal was the construction of a model has higher performance compared to the other works. We show that using preprocessing techniques in conjunction with ensemble classification methods may enhance the performance results. The combination of our preprocessing with ensembles and applied optimization algorithm could receive significant improvement than other works. It is the reason that the proposed model outperforms the benchmarks.

Conclusion

People purchase products on the Internet and give their reviews about them every second. These reviews affect the financial statements in companies noticeably. With the explosion of information on the Internet, it is difficult for ordinary people to make decisions about products. Sentiment Classification is a significant field in text mining that can help companies. The proposed model investigates meta-classifiers to increase the performance of the classification for Sentiment Classification on the Twitter datasets. We investigated the effects of the combination of sampling technique, PSO algorithm, and ensemble method on the classification performance. We characterized two weighting mechanisms and n-grams. The goal of our article was the suggestion of the effective model to increase the classification performance for classification tasks. PSO applied as a simple technique for solving optimization problems, which implies well-known ease of implementation and its convergence speed. Sampling as a particular technique is the superior approach; the bootstrapping method used and parameters of SVM optimized to improve the models. Boosting ensemble method in conjunction with SVM employed as base classifiers. The main advantage of the proposed comprehensive model is applying sampling techniques and PSO algorithm in preprocessing steps. We demonstrated the robustness of our model on the datasets. It appears that the IDF mechanism, bigram features, and the combination of them via bootstrapping sampling and PSO reaches the highest performance on the TSA datasets. We conclude that our investigation can be applicable for different social media analyses in the proposition of using ensemble technique, PSO, and bootstrapping the performance for Sentiment Classification. As future work, we are going to study the effect of other optimization algorithms and sampling techniques in this context.

Author Contributions

R. Asgarnezhad designed the experiments, carried out the data analysis, interpreted the results and wrote the manuscript. S. A. Monadjemi corrected the proofing the article. Soltanaghaei supported the article.

Acknowledgment

We thank the editor and all anonymous reviewers.

Conflict of Interest

The authors declare that they have no conflict of interest.

Abbreviations

TFIDF	Term Frequency-Inverse Document Frequency
TF	The frequency of word *t* in document *d*
N	The number of documents
Ft	The number of documents including word *t*
V_{ij}	The velocity of particles
x_{ij}	A group of particles
c_1	Cognitive parameter
c_2	social scaling parameter
MAE	Mean Absolute Error
RAE	Absolute Error

References

[1] E. Kouloumpis, T. Wilson, J.D. Moore, "Twitter sentiment analysis: The good the bad and the omg!," in Proc. Fifth International AAAI conf. on weblogs and social media: 538-541, 2011.

[2] F.H. Khan, S. Bashir, U. Qamar, "TOM: Twitter opinion mining framework using hybrid classification scheme," Decision Support Systems, 57: 245-257, 2014.

[3] N.F. Da Silva, E.R. Hruschka, E.R. Hruschka, "Tweet sentiment analysis with classifier ensembles," Decision Support Systems, 66: 170-179, 2014.

[4] A.C. Pandey, D.S. Rajpoot, M. Saraswat, "Twitter sentiment analysis using hybrid cuckoo search method," Information Processing & Management, 53: 764-779, 2017.

[5] H. Saif, M. Fernández, Y. He, H. Alani, "On stopwords, filtering and data sparsity for sentiment analysis of twitter," in Proc. Ninth International Conf. on Language Resources and Evaluation: 810–817, 2014.

[6] D. Tang, F. Wei, N. Yang, M. Zhou, T. Liu, B. Qin, "Learning Sentiment-Specific Word Embedding for Twitter Sentiment Classification," in Proc. The 52nd Annual Meeting of the Association for Computational Linguistics: 1555-1565, 2014.

[7] B. Besbinar, D. Sarigiannis, P. Smeros, "Tweet Sentiment Classification," Lausanne, 2014.

[8] A. Montejo-Ráez, E. Martínez-Cámara, M. T. Martín-Valdivia, L. A. Ureña-López, "Ranked wordnet graph for sentiment polarity classification in twitter," Computer Speech & Language, 28: 93-107, 2014.

[9] D.-T. Vo, Y. Zhang, "Target-Dependent Twitter Sentiment Classification with Rich Automatic Features," in Proc. IJCAI: 1347-1353, 2015.

[10] A. Go, R. Bhayani, L. Huang, "Twitter sentiment classification using distant supervision," CS224N Project Report, Stanford, 1: 1-6, 2009.

[11] L. Jiang, M. Yu, M. Zhou, X. Liu, T. Zhao, "Target-dependent twitter sentiment classification," in Proceedings of the 49th

Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-1: 151-160, 2011.
[12] A. Tripathy, A. Agrawal, S.K. Rath, "Classification of sentiment reviews using n-gram machine learning approach," Expert Systems with Applications, 57: 117-126, 2016.
[13] A.K. Tripathi, K. Sharma, M. Bala, "Parallel hybrid bbo search method for twitter sentiment analysis of large scale datasets using mapreduce," International Journal of Information Security and Privacy (IJISP), 13: 106-122, 2019.
[14] H. Saif, Y. He, H. Alani, "Alleviating data sparsity for twitter sentiment analysis," in Proc. the 21st International Conference on theWorld Wide Web: 2–9, 2012.
[15] L. Chen, W. Wang, M. Nagarajan, S. Wang, A. P. Sheth, "Extracting Diverse Sentiment Expressions with Target-Dependent Polarity from Twitter," ICWSM, 2: 50-57, 2012.
[16] R. Asgarnezhad, K. Mohebbi, "A Comparative Classification of Approaches and Applications in Opinion Mining," International Academic Journal of Science and Engineering, 2(1): 68-80, 2015.
[17] S. Monadjemi, R. Asgarnezhad, M. Soltanaghaei, "A High-Performance Model based on Ensembles for Twitter Sentiment Classification," Journal of Electrical and Computer Engineering Innovations (JECEI), 8(1): 41-52, 2020.
[18] R. Asgarnezhad, S.A. Monadjemi, M. Soltanaghaei, " FAHPBEP: A fuzzy Analytic Hierarchy Process framework in text classification," accepted in Majlesi Journal of Electrical Engineering, vol. 14, no. 3, 2020.
[19] A.K. Tripathi, K. Sharma, M. Bala, "Parallel hybrid bbo search method for twitter sentiment analysis of large scale datasets using mapreduce," International Journal of Information Security and Privacy (IJISP), 13: 106-122, 2019.
[20] S.H. Seyyedi, B. Minaei-Bidgoli, "Enhancing effectiveness of dimension reduction in text classification," International Journal on Artificial Intelligence Tools, 26(3): 1-21, 2017.
[21] S. Vashishtha, S. Susan, "Fuzzy rule based unsupervised sentiment analysis from social media posts," Expert Systems with Applications, 138: 1-15, 2019.
[22] R. Asgarnezhad, S.A. Monadjemi, M. Soltanaghaei, A. Bagheri, "SFT: A model for sentiment classification using supervised methods in Twitter," Journal of Theoretical & Applied Information Technology, 96(8): 2242-2251, 2018.
[23] A.K. Tripathi, K. Sharma, M. Bala, "Parallel hybrid bbo search method for twitter sentiment analysis of large scale datasets using mapreduce," International Journal of Information Security and Privacy (IJISP), 13: 106-122, 2019.
[24] A.K. Abbas, A. K. Salih, H. A. Hussein, Q.M. Hussein, S.A. Abdulwahhab, "Twitter Sentiment Analysis Using an Ensemble Majority Vote Classifier," Journal of Southwest Jiaotong University, 55: 1-7, 2020.
[25] N. Jiang, F. Tian, J. Li, X. Yuan, J. Zheng, "MAN: mutual attention neural networks model for aspect-level sentiment classification in SIoT," IEEE Internet of Things Journal, 7: 2901-2913, 2020.
[26] U. Naseem, I. Razzak, K. Musial, M. Imran, "Transformer based Deep Intelligent Contextual Embedding for Twitter sentiment analysis," Future Generation Computer Systems: 1-35, 2020.
[27] M.D. Samad, N.D. Khounviengxay, M.A. Witherow, "Effect of Text Processing Steps on Twitter Sentiment Classification using Word Embedding," arXiv preprint arXiv:2007.13027: 1-14, 2020.
[28] S. Sharma, A. Jain, "An Empirical Evaluation of Correlation Based Feature Selection for Tweet Sentiment Classification," in Proc. Advances in Cybernetics, Cognition, and Machine Learning for Communication Technologies, ed: Springer: 199-208, 2020.
[29] C.D. Manning, P. Raghavan, H. Schütze, Introduction to information retrieval vol. 1: Cambridge university press Cambridge, 2008.
[30] J.Han, M. Kamber. Data mining: concepts and techniques. Morgan Kaufmann Publishers–An Imprint of Elsevier, 500: 105-150, 2006.
[31] T.C. Hesterberg, "What teachers should know about the bootstrap: Resampling in the undergraduate statistics curriculum," The American Statistician, 69: 371-386, 2015.
[32] M.R. Chernick, W. González-Manteiga, R.M. Crujeiras, E.B. Barrios, Bootstrap methods. Springer, 2011.
[33] J.S. Haukoos, R.J. Lewis, "Advanced statistics: bootstrapping confidence intervals for statistics with “difficult” distributions," Academic emergency medicine, 12: 360-365, 2005.
[34] R. C. Eberhart, Y. Shi, J. Kennedy, Swarm intelligence: Elsevier, 2001.
[35] E. Fersini, A. Messina, F.A. Pozzi, "Sentiment Analysis: Bayesian Ensemble Learning," Decision Support Systems, 68: 26-38, 2014.

Biographies

Razieh Asgarnezhad received her B.Sc. and MSc. degrees in Computer Engineering from Kashan Azad University in 2009 and Arak Azad University in 2012, respectively. She is currently a Ph.D. candidate at the Department of Computer Engineering at Isfahan Azad University. Her current researches include Data Mining, Text Mining, Learning Automata, and Wireless Sensor Network.

S. Amirhassan Monadjemi is an Associate Professor at the University of Isfahan and Senior Lecturer, School of continuing and lifelong education at the National University of Singapore. He received a Ph.D. in Pattern Recognition from the University of Bristol in 2004. His research interests include Artificial Intelligence, Machine Vision, and Data Analysis.

Mohammadreza SoltanAghaei is an Associate Professor at the Islamic Azad University of Isfahan. He received a Ph.D. in Computer Network from UPM University of Malaysia in 2010. His research interests include Computer Networks and Data Mining.

How to cite this paper:
R. Asgarnezhad, A. Monadjemi, M. SoltanAghaei, “NSE-PSO: Toward an Effective Model using Optimization Algorithm and Sampling Methods for Text Classification,” Journal of Electrical and Computer Engineering Innovations, 8(2): 183-192, 2020.
DOI: 10.22061/JECEI.2020.7295.379
URL: http://jecei.sru.ac.ir/article_1460.html

Journal of Electrical and Computer Engineering Innovations
(JECEI)
Journal homepage: http://www.jecei.sru.ac.ir

Research paper

Parallel and Exact Method for Solving n-Similarity Problem

M. Mirhosseini, M. Fazlali*

Department of Data and Computer Science, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran.

Article Info

Article History:
Received 24 September 2019
Reviewed 20 November 2019
Revised 18 January 2020
Accepted 21 April 2020

Keywords:
n-Similarity
Parallel programming
Open-MP
Text document similarity

*Corresponding Author's Email Address:
fazlali@sbu.ac.ir

Abstract

Background and Objectives: n-similarity problem defined as measuring the similarity among $n \geq 3$ objects and finding a group of n objects from a dataset that have the most similarity to each other. This problem has been become an important issue in information retrieval and data mining. Theory of this concept is mathematically proven, but it practically has high memory complexity and is so time consuming. Besides, the solutions found by metaheuristics are not exact.
Methods: This paper is conducted to propose an exact method to solve n-similarity problem reducing the memory complexity and decreasing the execution time by parallelism using Open-MP. The experiments are performed on the application of text document resemblance.
Results: It has been shown that the memory complexity of the proposed method is decreased to $O(N^2)$, and the experimental results show that this method accelerates the speed of the computations about 5 times.
Conclusion: The simulated results of the proposed method display a good improvement in speed, the used memory space, and scalability compared with the previous exact method.

Introduction

The similarity is defined as the resemblance degree between two or more objects, phenomena or concepts. In various applications, designers want to build classes of objects based on the similarity criteria. So, it has an important role in automated classification, clustering methods, decision making, approximate reasoning, and diagnosis systems [1], [2], [3]. The similarity between two objects, known as 2-similarity, is mathematically defined as a function on domain U as $S_2: U \times U \longrightarrow [0,1]$ *or* $[-1,1]$. The value of 1 shows that the objects are identical, whereas 0 or -1 indicates that the objects are completely different. Various kinds of similarity measures including classic and fuzzy types have been suggested by researchers. Some well-known classic similarity/dissimilarity measures are Euclidean distances, Jaccard, Overlap, Dice, Pearson and Cosine coefficients that their effects are studied in data clustering [4].

Kaufman and Rousseeuw [4] compared several similarity measures on hierarchical clustering. Also, different types of fuzzy similarity measures have been developed. As an instance, a similarity measure between two fuzzy sets has been proposed in [6]. Its main idea is that if the intersection between two fuzzy sets is high, their similarity is identical. Two other fuzzy similarity measures based on the relative sigma count have been suggested in [7]. Using the Intuitionistic Fuzzy Sets (IFS) theories a fuzzy Cosine similarity measure and its weighted kind are proposed in [8]. These similarity measurements compute the similarity degree between two objects. But, in some applications, computing the similarity value among $n \geq 3$ objects or finding the most similar n-group among a dataset is required. The concept of n-similarity is mathematically defined by Keshavarzi et al. [1] at first. But, this method is infeasible due to the high time and memory complexities.

Therefore it is unable to handle the large datasets. To decrease the execution time of the algorithm and solving the problem in a reasonable time, a binary genetic algorithm has been exploited in [9]. Since the performance of metaheuristics varies in different problems, the effect of some other metaheuristics including, particle swarm optimization (PSO), gravitational search algorithm (GSA), imperialist competitive algorithm (ICA) and fuzzy imperialist competitive algorithm (FICA) has been studied and compared in solving the n- similarity problem and finding the most similar n-group of a dataset. The experiments show that the FICA has the best results [10].

Multi-core processors gain the market with increasing the number of cores per processor. But, the sequential programming model does not exploit multi-core systems well. Parallel programming techniques such as Open-MP present more effectively using of multiple processor cores to accelerate the speed of the algorithm [11]. Open-MP is utilized in various applications to reduce the execution time. For example, in [11], authors used Open-MP library to accelerate finding the maximum weighted clique. Also, the authors in [13] parallelized genetic algorithm to present a scalable method to solve large integer linear programming models derived from high-level synthesis of digital circuits. In [14] a new method for implementing the parallel breadth-first search algorithm for graph exploration on multi-core CPUs has been presented to accelerate the algorithm. In [15] the authors applied Open-MP to parallelize the quadrivalent quantum-inspired gravitational search algorithm in WSNs. This leads to improve the speedup faster than 4 times. Also, authors in [16] increased the speed of running the ant colony algorithm using Open-MP about 4.5 times. These researches encourage us to apply parallelism by multi-core CPU and Open-MP to improve the speed of our algorithm. Although metaheuristics [9], [10] can find a near optimal solution in a reasonable time, their results are not exact and they do not explore and compute the n-similarity value of all possible n-groups. Since Keshavarzi's method [1] is an exact method but suffers from high space complexity and its high execution time. Besides, the results of the metaheuristic methods [9], [10] are not exact. So, we are motivated to develop Keshavarzi's method as an exact one and improve its space complexity and decrease its execution time by parallelism using Open-MP. In fact, in the proposed approach, we perform some changes on recursive n-similarity relations introduced by Keshavarzi, et al. [1], and by that, the space complexity decreases from $O(N^n)$ to $O(N^2)$ for computing n-similarity for a dataset in the size of N. By this approach, it is not required to keep all similarity matrixes in the memory to be accessible in the recursive order. Moreover, since the proposed method involves several $nested\ for\ loops$, we apply the parallelism technique on the most time-consuming part of the algorithm to reduce the execution time. By this way, we would have an exact and executable method with lower space complexity that can run in a more reasonable time. This proposed method is better than Keshavarzi's [1] method from three aspects:

- The space complexity is low.
- Its execution time is less than the Keshavarzi's method.
- This method is scalable for larger datasets.

Also, it has two advantages in comparison with metaheurstics [9], [10] including:

- This method is an exact one.
- This method computes the n-similarity values for all possible groups of n objects.

The remainder of this study is organized as follows: the next section provides some preliminaries and related definitions. Then, the proposed parallel and exact method in solving the n-similarity problem is presented, and shows how this method can improve the space complexity. The experimental results as a case study on text document similarity are presented; and finally the paper is concluded in the last section.

Preliminaries

In this section some required definitions are presented from [1], [9].

Definition 1. Triangular norm (T-norm) is defined as a function $T: [0,1] \times [0,1] \longrightarrow [0,1]$ which satisfies the following conditions for all $x, y, w, z \in [0,1]$:

1. Commutatively: $T(x,y) = T(y,x)$,
2. Monotonicity: $T(x,y) \leq T(w,z), if\ x \leq w, and\ y \leq z$,
3. Associativity: $T(x, T(y,w)) = T(T(x,y), w)$,
4. Boundary : $T(x,0) = 0,\ T(x,1) = x$.

The minimum T-norm is a well-known instances of T-norms which has been applied in [1], [2] to n-similarity definition; $T_{min}(x,y) = \min(x,y)$.

Definition 2. The 2-similarity on domain U is a function $S: U \times U \longrightarrow [0,1]$ satisfying the following conditions:

1. Reflexivity: $for\ any\ x \in U, S(x,x) = 1$,
2. Symmetry: $for\ any\ x, y \in U, S(x,y) = S(y,x)$,
3. Transitivity: $for\ any\ x, y, z \in U, S(x,z) \geq S(x,y) \wedge S(y,z)$, where $\wedge$ is the minimum operator.

Definition 3. The 3-similarity on domain U is a function $S: U \times U \times U \longrightarrow [0,1]$ satisfying the following conditions:

1. Reflexivity: $for\ any\ x \in U, S(x,x,x) = 1$.
2. Symmetry: $for\ any\ x_1, x_2, x_3 \in U,\ S(x_1, x_2, x_3) = S(x_{i1}, x_{i2}, x_{i3})$ where i_1, i_2, i_3 is an arbitrary permutation of (1,2,3).
3. Transitivity property:

$for\ any\ t, x_1, x_2, x_3 \in U, S(x_1, x_2, x_3) \geq S(t, x_2, x_3) \wedge S(x_1, t, x_3) \wedge S(x_1, x_2, t)$, where $\wedge$ is the minimum T-norm.

The minimum T-norm is applied to generalize the 2-similarity to 3-similarity as (1). This equation satisfies all conditions of Definition 3.

$$S_3(x_1, x_2, x_3) = \min\big(S_2(x_1, x_2), S_2(x_1, x_3), S_2(x_2, x_3)\big) \quad (1)$$

Definition 4. $S: U \times U \times \ldots \times U \rightarrow [0,1]$ is the n–similarity function satisfying the following conditions:

1. Reflexivity: $for\ any\ x \in U, S(x, x, \ldots, x) = 1$,
2. Symmetry: $S(x_1, x_2, \ldots, x_n) = S(x_{i1}, x_{i2}, \ldots, x_{in})$ for all permutations $(i_1, i_2, \ldots, i_n)$ of $(1,2, \ldots, n)$.
3. Transitivity: $for\ all\ x_1, x_2, \ldots, x_n, z \in U, S(x_1, x_2, \ldots, x_n) \geq \min\{S(z, x_2, \ldots, x_n), \ldots, S(x_1, x_2, \ldots, x_{n-1}, z)\}$.

It was shown in [1] that the n-similarity can be achieved from the (n-1)-similarity satisfying all conditions of Definition 4. If S_{n-1} is the $(n-1)$-similarity on U and $x_1, x_2, \ldots, x_n \in U$, then:

$$S_n(x_1, x_2, \ldots, x_n) = \min\begin{pmatrix} S_{n-1}(x_2, x_3, \ldots, x_n), S_{n-1}(x_1, x_3, \ldots, x_n) \\ , \ldots, S_n(x_1, x_2, \ldots, x_{n-1}) \end{pmatrix} \quad (2)$$

The pseudo codes of 3-similarity, 4-similarity and n-similarity are presented in Algorithms 1, 2 and 3, respectively [1]. In these algorithm computing the similarity for each value of n is dependant on the previous similarities including $(n-1)$, $(n-2), \ldots, 3\ and\ 2$. So, these matrixes are required to save and this leads to increase the space complexity of this method and the program is unable to run for datasets which have more than 100 objects. In addition for several *nested for* the execution time of these algorithms are high and increase more by growing the size of dataset. In the next section, it is shown how we handle these problems.

Algorithm 1. The pseudo codes of 3-similarity

Input: *2-sim* matrix
Output: Max, *3-sim matrix*
Max=0;
for *i=1* to size of dataset
  for *j=i+1* to size dataset
    for *k=j+1* to size of dataset
      3-sim(i, j, k)=min{2-sim(i ,j),2-sim(i ,k),2-sim(j ,k)};
      If(3-sim(i, j, k)>Max)
        Max=3-sim(i, j, k);
end of for *i, j* and *k*
End

Algorithm 2. The pseudo codes of 4-similarity

Input: *3-sim* matrix
Output: Max, *3-sim matrix*
for *i=1* to size of dataset
  for *j=i+1* to size dataset
    for *k=j+1* to size of dataset
      for l=k+1 to size of dataset
4-sim(i, j, k, l)=min{3-sim(i, j, k),3-sim(i, j, l),3-sim(i, k ,l),3-sim(j, k, l)};
If(4-sim(i, j, k, l)>Max)
Max=4-sim(i, j, k, l);
end of for *i, j, k* and *l*
End

Algorithm 3. The pseudo codes of n-similarity

Input: *(n-1)-sim* matrix
Output: Max, *n-sim* matrix
for i_1=1 to size of dataset
  for i_2=i_1+1 to size of dataset
    .
    .
    for i_n=$i_{(n-1)}$ to size of dataset
      n-sim(i_1,i_2,…,i_n)=min{(n-1)-sim(i_2,i_3,…,i_n),(n-1)-sim(i_1,i_3,…,i_n),…, (n-1)-sim(i_1,i_2,…,i_{n-1})};
      If(n-sim(i, j, k)>Max)
        Max=n-sim(i, j, k);
end of for i_1 to i_n
End

The Proposed method

Regarding Algorithms 1, 2 and 3 proposed in [1], it can be shown that their executing time and the space complexity are $O(N^n)$, and the execution time and required memory space are grown by increasing the size of dataset (N) and n. Therefore, using this method for large datasets is practically impossible. Let N be the size of dataset. For computing the n-similarity, it is needed to calculate and store the $(n-1)$-similarity matrix in the size of N^{n-1}; as the same way, it is needed to calculate and store the $(n-2)$-similarity matrix in the size of N^{n-2} and so on. Therefore, the space complexity for computing the n-similarity would be equal to $N^n + N^{n-1} + N^{n-2} + \cdots + N^3 + N^2 = O(N^n)$. By the following proposed method, the space complexity would be decreased to $O(N^2)$. Let S_2, S_3, S_4, S_n be the 2-similarity, 3-similarity, 4-similarity and n-similarity respectively. We know:

$$S_4(x_1, x_2, x_3, x_4) = \min\begin{pmatrix} S_3(x_1, x_2, x_3), S_3(x_1, x_2, x_4), S_3(x_1, x_3, x_4), \\ S_3(x_2, x_3, x_4) \end{pmatrix} \quad (3)$$

By replacing S_3 from (1) to S_4 in (3), we have S_4 in terms of S_2 as (4).

$$S_4(x_1, x_2, x_3, x_4) = \min\begin{pmatrix} S_2(x_1, x_2), S_2(x_1, x_3), S_2(x_1, x_4), S_2(x_2, x_3), \\ S_2(x_2, x_4), S_2(x_3, x_4) \end{pmatrix} \quad (4)$$

In a similar way, S_n would be achieved as (5) by replacing S_{n-1} in terms of S_2. So, S_n can be computed regardless of previous $(n-1)$ similarity matrices and it just needs 2-similarity matrix as input instead of 2-similarity, 3-similarity,…,$(n-1)$-similarity matrixes. By this way, the space complexity is decreased from $O(N^n)$ to $O(N^2)$.

$$S_n(x_1, x_2, \ldots, x_n) = \min\big(S_2(x_1, x_2), S_2(x_1, x_3), \ldots, S_2(x_{n-1}, x_n)\big). \quad (5)$$

Besides, to decrease the execution time of the algorithms, we use parallelism technique using Open-MP. Open-MP is a portable implementation of parallel programs for shared memory multiprocessors. Recently, many algorithms are implemented in parallel using Open-MP achieving good speedup in addition to its simple implementation compared to other parallel methods like CUDA applied on GPU [17]. In this problem we have *nested for* loops. Either for loop can be run in parallel but, we make outer loop parallel to reduce number of forks and joins. Also, each thread gets its own private copy of variables j, k, etc. Also, since the iteration number of inner loops gradually is decreased, the dynamic scheduling is exploited, in which only some iterations of loop are allocated to threads and remaining iterations allocated to threads that complete their assigned iterations. Algorithms 4, 5 and 6 present the proposed method for computing the 3-similarity, 4-similarity and n-similarity respectively. In all these algorithms, we have n nested for loop for computing n-similarity. Although the time complexity does not change compared with [1], the execution time is improved using parallelizing. Instruction *#pragma omp parallel for* causes to divide the index of for i among running threads and each thread works on a part of outer loop. As an example if the size of dataset is 1000, and we have 4 threads, 250 tasks can be assigned to each thread. Each thread computes the related n-similarities and finds their maximum and stores the related maximum in *MaxTread[tid]* cell, where *tid* is the index of the thread. Finally, the maximum of array *MaxTread* is computed and stored as output *MAX*. Also, it is observable that all these pseudo codes just need the 2-similarity matrixes composing of the similarity between all possible pair wise objects.

This matrix is a symmetric one with the size of $N \times N$ that N is the size of dataset. The elements on the main diagonal are one; and to avoid producing results with frequent objects, the elements below the main diagonal are set to zero. So, each inner for loop starts from the next value of previous loop index.

Algorithm 4. The pseudo codes of the proposed 3-similarity

```
Input: 2-sim matrix
Output: Max, 3-sim of each group
#pragma omp parallel for
 for i=1 to size of dataset
  for j=i+1 to size dataset
    for k=j+1 to size of dataset
      3-sim=min{2-sim(i , j),2-sim(i ,k),2-sim(j ,k)};
      Write:3-sim
        If(3-sim>MaxTread[tid])
          MaxTread[tid]=3-sim;
 end of for i, j and k
MAX=maximum of MaxTread
End
```

Algorithm 5. The pseudo codes of the proposed 4-similarity

```
Input: 2-sim matrix
Output: Max;
#pragma omp parallel for
for i=1 to size of dataset
  for j=i+1 to size dataset
    for k=j+1 to size of dataset
      for l=k+1 to size of dataset
 begin
4-sim=min{2-sim(i, j),2-sim(i, k),2-sim(i, l),2-sim(j, k),2-sim(j,l),2-sim(k,l)};
Write:4-sim
If(4-sim> MaxTread[tid])
          MaxTread[tid]=4-sim;
 end of for i, j, k and l
MAX=maximum of MaxTread
End
```

Algorithm 6. The pseudo codes of the proposed n-similarity

```
Input: 2-sim matrix
Output: Max
#pragma omp parallel for
 for i1=1 to size of dataset
  for i2=i1+1 to size of dataset
    .
      .
     for in=i(n-1) to size of dataset
       n-sim=min{2-sim(i1,i2),2- sim(i1,i3),..., 2-sim(in-1,in)};
       Write:n-sim
        If(n-sim> MaxTread[tid])
          MaxTread[tid]=n-sim;
end of for i1 to in
MAX=maximum of MaxTread
End
```

Results and Discussion

To compare the proposed method with the exact method suggested in [1], some experiments are conducted. One of the applications of n-similarity is in text similarity. The purpose is to find n documents among N texual documents in such away they have the most similarity to each other. Also the similarity among all possible n permutations is computed. The Re0 dataset from the Reuters repository [18] are used in this work, contained 1504 newspaper articles, 2886 key words and 31 classes. To work with the textual datasets some preprocessing steps are required to do. This procedure performs as follows and the output is the 2-similarity matrix which is given as input to the n-similarity algorithms.

Step 1: The first is tokenization in which, all digits and symbols are removed and the strings occurred before $space$, . , :, ; , − , ? , ! , (,) [,] etc. are extracted and considered as a word.

Step 2: in this step, all stop-words like *a, the, is, are,* etc. are eliminated using a pre-supplied list named Weka machine learning workbench [19] included 527 stop-words.

Step 3: There are some words with similar theme and different morphological concept. These words would be

mapped into their stem to treat as a single word. For instance both words *computing* and *computation* are maped into the stem *comput*. The Porter's suffix-stripping algorithm [20] is applied for stemming the words.

Step 4: In the weighting step, each extracted term gets a numerical weight. $T = \{t_1, t_2, \dots, t_l\}$ is all occurred words in the dataset and $d_i = \{w_{1i}, w_{2i}, \dots, w_{li}\}$ is the feature vector of document d_i, where w_{ki} is the weight of word t_k in document d_i. The weighting process is done by the TFIDF as (6).

$$w_{ki} = TFIDF(d_i, t_k) = TF(d_i, t_k) \times \log(\frac{|D|}{DF(t_k)}), \quad (6)$$

In this relation, *D* is the size of the dataset; $TF(d_i, t_k)$ is the frequency of term t_k in the document d_i and $DF(t_k)$ is the number of documents in which term t_k occurs in [21].

Step 5: In this step, the terms with the weight less than a predefined threshold are removed.

Step 6: In this step, to compose the 2-Similarity matrix, the similarity between all documents is computed by Cosine similarity measurement, which is a usual measurement for text document applications. Equation (7) shows this similarity measurement which computes the similarity degree between documents d_i and d_j using their feature vectors as $d_i = \{w_{1i}, w_{2i}, \dots, w_{li}\}$ [21].

$$Cosine - Sim(d_i, d_j) = \frac{\sum_{k=1}^{n} w_{ki} w_{kj}}{\sqrt{\sum_{k=1}^{n} w_{ki}^2 \sum_{k=1}^{n} w_{kj}^2}} \quad (7)$$

Finally, the 2-similarity matrix give as input to Algorithms 4, 5 or 6 to compute the 3-similarity, 4-similarity or n-similarity respectively. The simulations are performed on a laptop with CPU: Intel(R) Core(TM) i7-4702 MQ CPU @ 2.20GHz, 4 GB RAM. The results are averaged over 5 running of the algorithms. Fig. 1, Fig. 2 and Fig. 3 respectively show the results for 3-similarity, 4-similarity and 5-similarity. Each table presents the speedup of using the proposed method for different size of Re0 dataset from 200 to 1504 by changing the number of cores from 2 to 8 (showing by $S2$ to $S8$). In Fig. 1, for dataset with the size of 200, the speedup is enhanced from 1.45 to 4.29 by changing the number of cores from 2 to 8. Also, for dataset in the size of 1,504 the speedup is enhanced from 1.89 to 5.43 in average. The averaged speedup for different size of dataset is increased from 1.66 to 5 by changing the number of cores. Fig. 2 shows the results in the case of 4-similarity problem. As an example, for dataset with 400 documents, the speedup is gradually enhanced from 1.47 to 5.21 for different number of threads. Also, in the case of dataset with the size of 1,504, the speedup is 1.73 using 2 cores and is 5.68 for 8 cores. The averaged speedup for different size of dataset varies from 1.68 to 5.38 using 2 cores and 8 cores respectively. Similarly, Fig. 3 demonstrates the results for 5-similarity problem. The averaged speedup increased from 1.61 to 4.82 using different number of cores starting from 2 to 8. Regarding these figures, the speedup of 5, 5.38 and 4.82 are averagely achieved for 3-similariy, 4-similarity and 5-similarity problems by 8 cores. Fig. 1, Fig. 2 and Fig. 3 show that the speedup is generally improved by increasing the size of dataset and increasing the number of threads for 3-similarity, 4-similarity and 5-similarity problems. In fact, for larger datasets, increasing the number of threads leads to more speedup.

Because in the case of smaller datasets dividing the data on more number of cores has more overhead in comparison with dividing large datasets on a same number of cores. Fig. 4 represents the efficiency using 8 cores for 3, 4 and 5-smilarity problems for different size of datasets. It can be seen that for 3-similarity problem, the efficiency is increased from 53.6% for a dataset with the size of 200 to 67.8% for the dataset with the size of 1,504. In addition, for 4-similarity problem, the efficiency of using 8 cores changes as 57.7% to 71.0% by changing the size of dataset from 200 to 1,504. Also, the efficiency of using 8 cores is computed as 51.3% for a dataset with the size of 200 and increased to 67.3% for dataset with 1,504 documents. It is observable that by growing the size of dataset, the efficiency is increased.

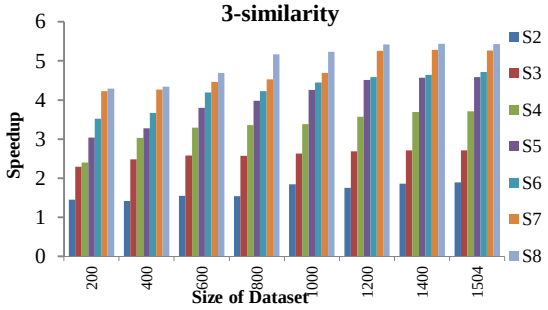


Fig. 1: The speedup of the proposed method in solving the 3-similarity for different number of threads and different size of datasets.

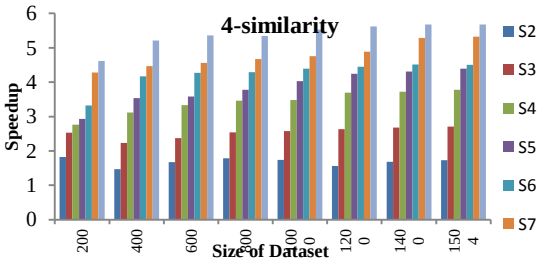


Fig. 2: The speedup of the proposed method in solving the 4-similarity for different number of threads and different size of datasets.

Table 1 presents the execution time of the Keshavarzi's method [1] and our proposed method in sequential and parallel with 8 threads for 3-similarity, 4-similarity and 5-similarity by changing the size of dataset. In this table "-" means that the method is unable to solve the problem.

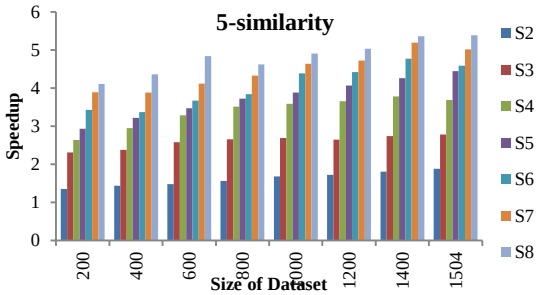


Fig. 3: The speedup of the proposed method in solving the 5-similarity for different number of threads and different size of datasets.

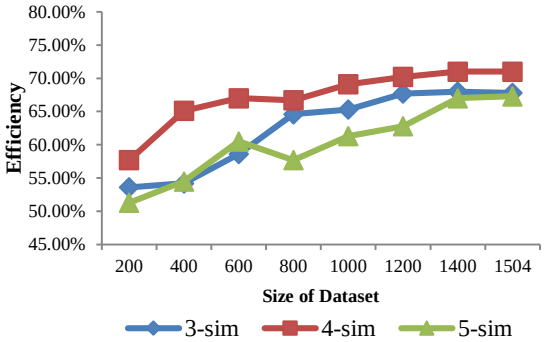


Fig. 4: The of efficiency using 8 threads for different size of dataset for 3, 4 and 5-similarity problems.

The columns entitled by "[1]" shows that although the Keshavarzi's method [1] is theatrically true, it is practically unable to solve the n-similarity problem for $n \geq 5$. Besides, for $n = 4$ the proposed method in [1] just can work on datasets with the size of less than 200. For the 3-similarity problem, the Keshavarzi's method [1] is applicable for datasets which are contained less than 1400 documents. These are due to the high memory space which this method needs to keep matrixes to be accessible in recursive order. While, our proposed method is applicable for all examined size of datasets. Because, it just needs to keep the 2-similarity matrix in the size of N^2 to solve the 3, 4, 5 and generally n-similarity problems. Moreover, in this table, it clearly illustrates the decreasing execution time using parallelism in comparison with sequential version of the proposed method. So, our proposed method is a scalable version of exact method proposed in [1]. Table 2 shows the approximate required space memory to keep the matrixes in Keshavarzi's method [1] and our suggested approach. It can be clearly seen that the proposed method needs very low memory in comparison with [1].

Besides, the used space memory of the proposed method is constant for 3, 4 and 5-similarity problems. This table truly shows why the proposed method in [1] is unable to work for large datasets and bigger amount of n as have been specified by "-" at Table 1.

Table 1: Comparison of the expectation time (sec.) of the proposed method in sequential and parallel for 3, 4 and 5-similarity

Dataset size	$n = 3$			$n = 4$			$n = 5$		
	[1]	Seq.	Par.	[1]	Seq.	Par.	[1]	Seq.	Par.
200	0.032	0.025	0.005	1.76	1.669	0.320	-	211.652	51.496
400	0.31	0.252	0.058	-	25.268	4.849	-	2899.12	664.935
600	0.882	0.868	0.185	-	128.207	23.919	-	4576.58	945.574
800	1.741	1.698	0.328	-	406.095	76.047	-	9543.86	2065.771
1000	3.621	3.570	0.682	-	991.948	179.375	-	14763.76	3006.875
1200	6.39	6.315	1.165	-	2057.19	366.048	-	18673.91	3712.506
1400	9.024	8.773	1.612	-	3824.16	673.267	-	24983.13	4661.031
1504	-	10.38	1.912	-	5262.34	926.468	-	26459.87	4909.066

Table 2: Comparison of approximate size of required memory for the proposed method and Keshavarzi's method [1] in solving n-similarity for $n = 3, 4$ and 5

Dataset size	$n = 3$		$n = 4$		$n = 5$	
	[1]	Proposed method	[1]	Proposed method	[1]	Proposed method
200	8 MB	**40 KB**	1.6 GB	**40 KB**	320 GB	**40 KB**
400	64 MB	**160 KB**	25.6 GB	**160 KB**	10.2 TB	**160 KB**
600	216 MB	**360 KB**	129.8 GB	**360 KB**	77.8 TB	**360 KB**
800	512 MB	**640 KB**	410.1 GB	**640 KB**	328 TB	**640 KB**
1000	1 GB	**1 MB**	1 TB	**1 MB**	1 PB	**1 MB**
1200	1.7 GB	**1.4 MB**	2 TB	**1.4 MB**	2.4 PB	**1.4 MB**
1400	2.7 GB	**2 MB**	3.8 TB	**2 MB**	5.3 PB	**2 MB**
1504	3.4 GB	**2.2 MB**	5.1 TB	**2.2 MB**	7.7 PB	**2.2 MB**

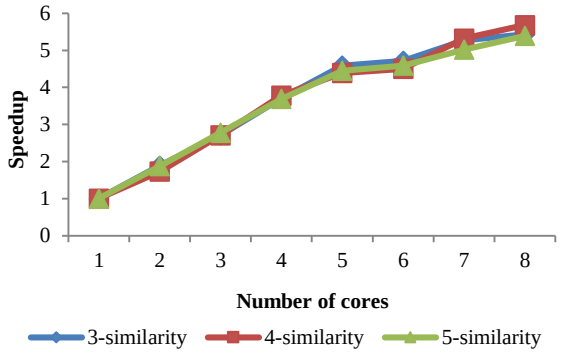


Fig. 5: The speedup of proposed method for Re0 dataset in the size of 1504 by changing the number of cores.

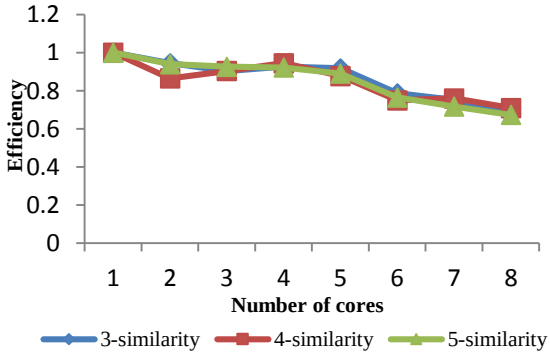


Fig. 6: The efficiency of the proposed method for Re0 dataset in the size of 1504 by changing the number of cores.

Fig. 5 shows the speedup by changing the number of cores from 2 to 8 for Re0 dataset with the size of 1,504, for 3-similarriy, 4-similarity and 5-similarity problems. For the 3-similrity, the speedup is changed from 1 to 5.43. For 4-similarity it is reached from 1 to 5.68 and in the case of 5-similarity, the speedup increases from 1 to 5.39 by changing the number of threads from 1 to 8.

It is observable that the trend of speedup for these three problems is so close to each other. Also, Fig. 6 shows the efficiency for Re0 dataset in the size of 1,504 for different number of cores from 1 to 8, for 3-similarity, 4-similarity and 5-similarity problems. The efficiency respectively reaches to 67.8%, 71.0% and 67.3% for 3-similarity, 4-similarity and 5-similarity by an 8-core CPU. The trend of this figure shows decreasing the efficiency from 100% to 68% in average. Also, as in Fig. 5, the efficiency trend is for the three problems is similar. Generally, the proposed parallel method leads to decreasing the execution time and increasing the speedup especially for larger datasets. As a consequence our method is able to deliver the exact result in a practically feasible time in comparison with the state-of-the-art exact method.

Conclusion

In this study, the theory of the n-similarity problem is reviewed. The executing time and the space complexity of the previous exact method to solve this problem are high and increase more by growing the size of dataset. In addition, the metaheuristics proposed to solve this problem generate non-deterministic solutions and are unable to compute the n-similarity values for all possible groups of n objects. Therefore, in this paper we proposed an exact method with low space complexity and improve its running time using parallelism by Open-MP. The experiments performed on a textual dataset by varying its size and reported results show that the proposed parallel method averagely accelerate the speed of the method as 5, 5.38 and 4.82 times for 3-similariy, 4-similarity and 5-similarity problems respectively. Also, it was mathematically proved that for each value of n, the space complexity was improved to $O(N^2)$, in which N is the size of dataset.

Author Contributions

M. Mirhosseini carried out the experiment, and wrote the manuscript. M. Fazlali helped supervise and write the paper.

Acknowledgment

The authors gratefully express sincere thanks to dear reviewers and editors for their guidance and valuable comments.

Conflict of Interest

The author declares that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancy have been completely observed by the authors.

Abbreviations

(n-1)-sim	(n-1)-similarity
d_i	Document i
w_{ki}	the weight of word t_k in document d_i
2-sim	2-similarity
3-sim	3-similarity
Cosine_Sim	Cosine similarity measurement
MAX	maximum
MaxTread	Found maximum by each thread
min	minimum
N	Size of dataset
n-Sim	n-similarity
S	similarity
S_2	2-similarity
S_3	3-similarity
S_4	4-similarity
S_n	n-similarity
T	The set of terms
TF	Term frequency
tid	Thread ID
DF	number of documents

References

[1] M. Keshavarzi, M. A. Dehghan, M. Mashinchi, "Applications of classification based on similarities and dissimilarities," Fuzzy Information and Engineering, 4(1): 75-92, 2012.

[2] M. Keshavarzi, M. A. Dehghan, M. Mashinchi, "Classification based on 3-similarity, Iranian Journal of Mathematical Sciences and Informatics," 6(1): 7-21, 2011.

[3] M. Keshavarzi, "Classification based on similarity and dissimilarity", PhD thesis, Shahid Bahonar University of Kerman, Iran, 2010.

[4] S. Theodoridis, K. Koutroumbas, Pattern recognition, Academic Press, 2003.

[5] L. Kaufman, P. J. Rousseeuw, Finding Group in Data An Introduction to Cluster Analysis, Wiley, New York, 2005.

[6] W. J. Wang, "New similarity measure on fuzzy sets and on elements", Fuzzy Sets and Systems, 85(3): 305-309, 1997.

[7] H. Rezaei, M. Emoto, M. Mukaidono, "New similarity measure between two fuzzy sets," Journal of Advanced Computational Intelligence and Intelligent Informatics, 10(6): 946-953, 2006.

[8] J. Ye, "Cosine similarity measures for intuitionistic fuzzy sets and their applications," Mathematical and Computer Modeling, 53: 91–97, 2011.

[9] M. Mirhoseini, M. Mashinchi, H. Nezamabadi-pour, "Improving n-Similarity problem by genetic algorithm and its application in text document resemblance," Fuzzy Information and Engineering, 6: 263-278, 2014.

[10] M. Mirhoseini, H. Nezamabadi-pour, "Metaheuristic Search Algorithms in Solving the n-Similarity Problem," Fundamenta Informaticae, 152(2): 145-166, 2017.

[11] K. Lakshmanan, S. Kato, R. Rajkumar, "Scheduling Parallel Real-Time Tasks on Multi-core Processors," in Proc. 2010 31st IEEE Real-Time Systems Symposium: 259-268, 2010.

[12] M. K. Fallah, V. S. Keshvari, M. Fazlali, "A Parallel Hybrid Genetic Algorithm for Solving the Maximum Clique Problem," in Proc. High-Performance Computing and Big Data Analysis. TopHPC 2019. Communications in Computer and Information Science, 891: 378-393, 2019.

[13] M. K. Fallah, M. Mirhosseini, M. Fazlali, M. Daneshtalab, "Scalable Parallel Genetic Algorithm For Solving Large Integer Linear Programming Models Derived From Behavioral Synthesis," in Proc. 28th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP): 390-394, 2020.

[14] S. Hong, T. Oguntebi, K. Olukotun, "Efficient Parallel Graph Exploration on Multi-Core CPU and GPU", 2011 International Conference on Parallel Architectures and Compilation Techniques, Galveston, TX, pp. 78-88, 2011.

[15] M. Mirhosseini, M. Fazlali, G. Gaydadjiev, "A Parallel and Improved Quadrivalent Quantum-Inspired Gravitational Search Algorithm in Optimal Design of WSNs," High-Performance Computing and Big Data Analysis. TopHPC 2019. Communications in Computer and Information Science, 891: 352-366, 2019.

[16] P. Delisle, M. Krajecki, M. Gravel, C. Gagné, "Parallel implementation of an ant colony optimization metaheuristic with OpenMP," In Proceedings of the 3rd European Workshop on OpenMP (EWOMP'01): 1-7, 2001.

[17] L. Dagum, M. Menon. "OpenMP: an industry standard API for shared-memory programming," IEEE Computational Science and Engineering, 5(1): 46-55, 1998

[18] http://www.daviddlewis.com/resources/testcollections/reuters21578/.

[19] http://www.ai.mit.edu/projects/jmlr/papers/volume5/lewis04a/a11-smart-stop-list/.

[20] M. F. Porter, "An algorithm for suffix stripping", Program, 14(3): 130–137, 1980.

[21] C. Qimin, G. Qiao, W. Yongliang, W. Xianghu, "Text clustering using VSM with feature clusters", Neural Computing and Applications, vol 26, pp. 995- 1003, 2015.

Biographies

Mina Mirhosseini received her B.Sc. and M.Sc. degrees from Shahid Bahonar University, Kerman, Iran, in Computer Engineering and Computer Science, respectively. Currently, she is working towards a Ph.D. degree in Computer Science at Shahid Beheshti University, Tehran, Iran. Her research interests include parallel processing, metaheuristic, text processing and wireless sensor networks.

Mahmood Fazlali received B. Sc in computer engineering from Shahid Beheshti University (SBU) in 2001. Then he received M.Sc from University of Isfahan in 2004, and PhD from SBU in 2010 in computer architecture. He performed researches on reconfigurable computing systems in computer engineering lab at Technical University of Delft (TUDelft) as a postdoc researcher. Now, he is working as assistant professor at department data and computer sciences at SBU. His research interest includes parallel processing, reconfigurable computing, and data sciences.

How to cite this paper:
M. Mirhosseini, M. Fazlali, "Parallel and Exact Method for Solving n-Similarity Problem," Journal of Electrical and Computer Engineering Innovations, 8(2): 193-200, 2020.

DOI: 10.22061/JECEI.2020.7247.377

URL: http://jecei.sru.ac.ir/article_1461.html

Journal of Electrical and Computer Engineering Innovations
(JECEI)
Journal homepage: http://www.jecei.sru.ac.ir

Research paper

Coordinated Model Predictive DC-Link Voltage, Current, and Electromagnetic Torque Control of Wind Turbine with DFIG under Grid Faults

Z. Dehghani Arani[1], S.A. Taher[1,], M.H. Karimi[1,2], M. Rahimi[1]*

[1]*Department of Electrical Engineering, University of Kashan, Kashan, Iran.*

[2]*Iranian Oil Pipeline and Telecommunication Company, Iran.*

Article Info

Article History:
Received 23 October 2019
Reviewed 11 December 2019
Revised 01 January 2020
Accepted 02 March 2020

Keywords:
Coordinated model predictive control
DFIG-based WT
Grid code requirements
Grid faults

*Corresponding Author's Email Address:
ataher@kashanu.ac.ir

Abstract

Background and Objectives: The wind turbines (WTs) with doubly fed induction generator (DFIG) have active and reactive power as well as electromagnetic torque oscillations, rotor over-current and DC-link over-voltage problems under grid faults. Solutions for these problems presented in articles can be classified into three categories: hardware protection devices, software methods, and combination of hardware and software techniques.
Methods: Conventional protection devices used for fault ride through (FRT) capability improvement of grid-connected DFIG-based WTs impose difficulty in rotor side converter (RSC) controlling, causing failure to comply with grid code requirements. Hence, the main idea in this paper is to develop a novel coordinated model predictive control (MPC) for the power converters without need to use any auxiliary hardware. Control objectives are defined to maintain DC-link voltage, rotor current as well as electromagnetic torque within permissible limits under grid fault conditions by choosing the best switching state so as to meet and exceed FRT requirements. Model predictive current and electromagnetic torque control schemes are implemented in the RSC. Also, model predictive current and DC-link voltage control schemes are applied to grid side converter (GSC).
Results: To validate the proposed control method, simulation studies are compared to conventional proportional-plus-integral (PI) controllers and sliding mode control (SMC) with pulse-width modulation (PWM) switching algorithm. In different case studies comprising variable wind speeds, single-phase fault, DFIG parameters variations, and severe voltage dip, the rotor current and DC-link voltage are respectively restricted to 2 pu and 1.2 times of DC-link rated voltage by the proposed MPC-based approach. The maximum peak values of DC-link voltage are 1783, 1463 and 1190 V by using PI control, SMC and the proposed methods, respectively. The maximum peak values of rotor current obtained by PI control, SMC and the proposed strategies are 3.23, 3.3 and 1.95 pu, respectively. Also, PI control, SMC and the proposed MPC methods present 0.8, 0.4 and 0.14 pu, respectively as the maximum peak values of electromagnetic torque.
Conclusion: The proposed control schemes are able to effectively improve the FRT capability of grid-connected DFIG-based WTs and keep the values of DC-link voltage, rotor current and electromagnetic torque within the acceptable limits. Moreover, these schemes present fast dynamic behavior during grid fault conditions due to modulator-free capability of the MPC method.

Introduction

Due to development in power electronic technology, variable speed wind energy conversion systems (WECSs)

are integrated into power systems [1]. Among different types of WECSs, doubly fed induction generator (DFIG)-based wind turbines (WTs) are widely used because of flexible operation at different speeds, high energy transfer capability, requirement of low cost and small size power electronics system, control capability of active and reactive powers, etc. [2].

During faults and voltage dips occurred in a power grid, the stator current of DFIG-based WT increases as a result of the direct connection of stator windings to the power grid. Since there is magnetic coupling between stator and rotor windings, high rotor inrush currents and DC-link capacitor over-voltage are created [3]. In order to address the risk of damage to rotor side converter (RSC) and DC-link capacitor, and to improve the fault ride through (FRT) capability, researchers have proposed three different types of solutions: hardware protection techniques, software methods, and combination of hardware and software methods.

Crowbar is an old hardware protection device to maintain power electronics and to improve FRT behavior of DFIG-based WTs. It suggests disconnecting WT from the grid during severe electrical faults. Although this protection device avoids high rotor inrush currents and leads to RSC isolation, the DFIG-based WT operates similar to a squirrel cage induction generator drawing a considerable amount of reactive power from the power grid [4], [5]. On the other hand, recent grid codes require DFIG-based WTs to stay connected to the power grid during and after faults, and to provide FRT capability [6]. Therefore, developed control strategies have been proposed to improve the DFIG performance during the voltage dip and crowbar activation [7], [8].

A set of thyristor controlled resistors that are connected to the rotor windings have been presented in [3] in order to limit the high current and to provide a bypass for it in the rotor circuit. Furthermore, superconducting magnetic energy storage-fault current limiter (SMES-FCL) [9], dynamic voltage restorer (DVR) [10], static synchronous compensator (STATCOM) [11] and static volt ampere reactive compensator (SVC) [12] are employed as hardware protection methods to improve FRT behavior of DFIG-based WTs. In [13], the authors used a superconducting coil (SC) in the DFIG-based WT's DC-link. The proposed hardware solution acts as a FCL during severe faults of power systems to reduce the rotor and stator over-currents and also the DC-link over-voltage, while it acts as an energy storage device during normal operating conditions to smooth out output power fluctuations. FRT requirement of current-based protection devices which their accurate operation necessitates fault current provision by resources for a given period of time has been discussed in [14]. In [15], a modified DC chopper has been proposed not only to keep the DC-link voltage in acceptable range, but also to limit the rotor transient over-current in a permissible level without requiring the high rated current antiparallel diodes in the RSC. A novel inductive superconducting fault current limiter (SFCL)-based protection strategy with demagnetization technology has been presented in [16] to enhance the FRT capability of DFIG-based WTs. However, hardware solutions are still expensive as well as, they impose additional maintenance costs.

In addition to hardware protection techniques, advanced control strategies have been introduced to limit the DC-link over-voltage and rotor over-current which might be categorized under software solutions [17]-[23], [25]-[29]. The authors in [21] implemented proportional-plus-integral (PI) controllers combined with Lyapunov-based nonlinear control method in order to enhance the transient behavior of DFIG-based WT. In this FRT method, the rotor back-EMF voltage compensation leads to limit inrush current of the rotor and electromagnetic torque oscillations. In [22] and [23], feed-forward transient control approach was utilized for RSC so as to enhance the FRT capability.

In recent years, model predictive control (MPC) has attracted lots of attention because it provides a free-modulation technique to control power converters, as well as simple and flexible control approach while constraints and nonlinearities are included. In [24], a model-based predictive controller has been presented for DFIG direct power control in RSC which the control law has been derived by optimization of the difference between the predicted active and reactive powers and their references. Model predictive current control has been proposed as FRT improvement strategy for DFIG-based WT in [25]-[27].

Nevertheless, none of these works utilize the benefits of MPC to control and limit DC-link voltage and electromagnetic torque during fault conditions. An application of fuzzy logic controller with type of Takagi–Sugeno–Kang tuned using adaptive neuro-fuzzy inference system (ANFIS) and MPC theory for power converters was designed in [27] to limit the DC-link over-voltage and rotor over-current under fault conditions, although improving transient electromagnetic torque has not been considered.

As the DFIG-based WT has nonlinear dynamics, nonlinear controllers have been applied to power converters in [28], [29]. The transient behavior of DFIG by using sliding mode controllers and two protection circuits of crowbar and DC-chopper under severe grid faults was discussed in [28]. Even though the proposed nonlinear control technique can regulate DC-link voltage within acceptable limit, it tends to fail in limiting the rotor current.

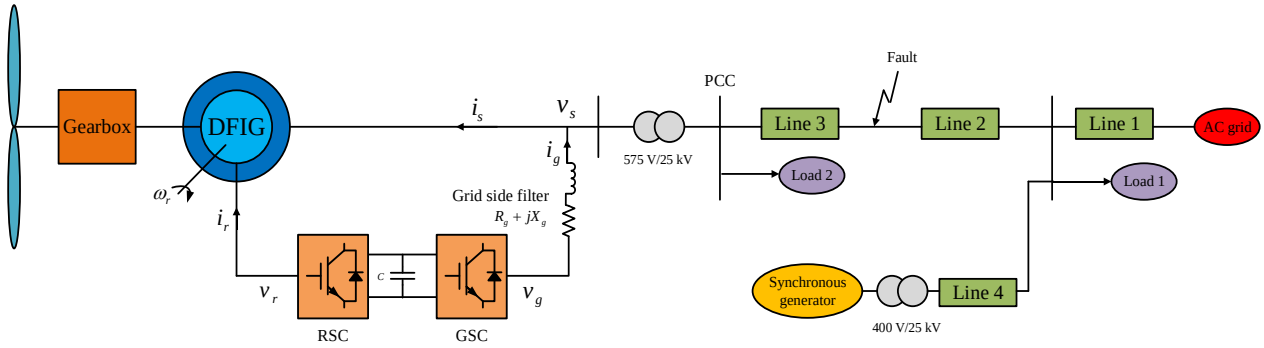


Fig. 1: The grid-connected WT with DFIG under study.

Because of high cost and low system reliability of using hardware solutions and unsatisfactory performance of some software solutions, proposing a new control approach as a software solution in order to enhance FRT capability of DFIG-based WT still seems indispensable; a novel approach which is robust against changes in the system model without requiring auxiliary protection devices. Since the PI controllers will be easily saturated under severe voltage dips, these controllers with pulse-width modulation (PWM) switching algorithm in both RSC and grid side converter (GSC) are substituting by MPC schemes in this paper. As the main contribution of this paper, predictive control of electromagnetic torque and DC-link voltage during fault condition are employed in coordinated MPC schemes of RSC and GSC, respectively. Moreover, power converter switching signals are obtained by the MPC theory without using any additional modulation techniques to improve the FRT capability of DFIG-based WTs during extreme voltage dips. And finally, the proposed control structure performance in the FRT capability enhancement of DFIG-based WTs is compared with sliding mode control (SMC) and PI control by several simulations in the MATLAB/SIMULINK environment.

The remaining parts of the paper are given as follows. The next section presents modeling and conventional control structure of DFIG-based WTs. Then, the SMC for FRT improvement, which has been proposed in [28], is reviewed. Next, the design of improved MPC-based control schemes for enhancing transient behavior of DFIG-based WT under grid fault condition is described. Simulation results are discussed and concluded in the next sections to validate the efficiency of the proposed control approach.

Modeling and Conventional Control of DFIG-Based WT

The schematic diagram of the DFIG-based WT test system is illustrated in Fig. 1. With respect to the indicated stator and rotor current directions, the DFIG-based WT dynamics in a synchronous reference frame is derived. The voltages and also fluxes of stator and rotor circuits in a dq reference frame rotating at angular speed of ω, in per unit values, and referred to the stator side are presented as follows [21], [30]:

$$v_{sdq} = R_s i_{sdq} + j\omega\psi_{sdq} + \frac{1}{\omega_b}\frac{d\psi_{sdq}}{dt} \tag{1}$$

$$v_{rdq} = R_r i_{rdq} + j\omega_2\psi_{rdq} + \frac{1}{\omega_b}\frac{d\psi_{rdq}}{dt} \tag{2}$$

$$\psi_{sdq} = L_s i_{sdq} + L_m i_{rdq} \tag{3}$$

$$\psi_{rdq} = L_r i_{rdq} + L_m i_{sdq} \tag{4}$$

In accordance with (1), the stator flux state space equations can be written as

$$\begin{cases} \dfrac{d\psi_{sd}}{dt} = \omega_b\left(v_{sd} - R_s i_{sd} + \omega\psi_{sq}\right) \\ \dfrac{d\psi_{sq}}{dt} = \omega_b\left(v_{sq} - R_s i_{sq} - \omega\psi_{sd}\right) \end{cases} \tag{5}$$

Using (2)-(4), the state space equations of rotor circuit are given as

$$\begin{cases} \dfrac{di_{rd}}{dt} = -\dfrac{\omega_b R_r'}{L_r'} i_{rd} + \omega_b\omega_2 i_{rq} - \dfrac{\omega_b}{L_r'} E_d + \dfrac{\omega_b}{L_r'} v_{rd} \\ \dfrac{di_{rq}}{dt} = -\dfrac{\omega_b R_r'}{L_r'} i_{rq} - \omega_b\omega_2 i_{rd} - \dfrac{\omega_b}{L_r'} E_q + \dfrac{\omega_b}{L_r'} v_{rq} \end{cases} \tag{6}$$

where R_r', L_r' and E_{dq} are given by

$$R_r' = R_r + \left(\frac{L_m}{L_s}\right)^2 R_s \tag{7}$$

$$L_r' = L_r - \frac{L_m^2}{L_s} \tag{8}$$

$$E_{dq} = \frac{L_m}{L_s}\left(v_{sdq} - j\omega_r\psi_{sdq} - \frac{R_s}{L_s}\psi_{sdq}\right) \tag{9}$$

The electromagnetic torque directly depends on stator flux and rotor current as

$$T_e = \frac{L_m}{L_s}\left(\psi_{sq} i_{rd} - \psi_{sd} i_{rq}\right) \tag{10}$$

The harmonic pollution of the output current in the GSC

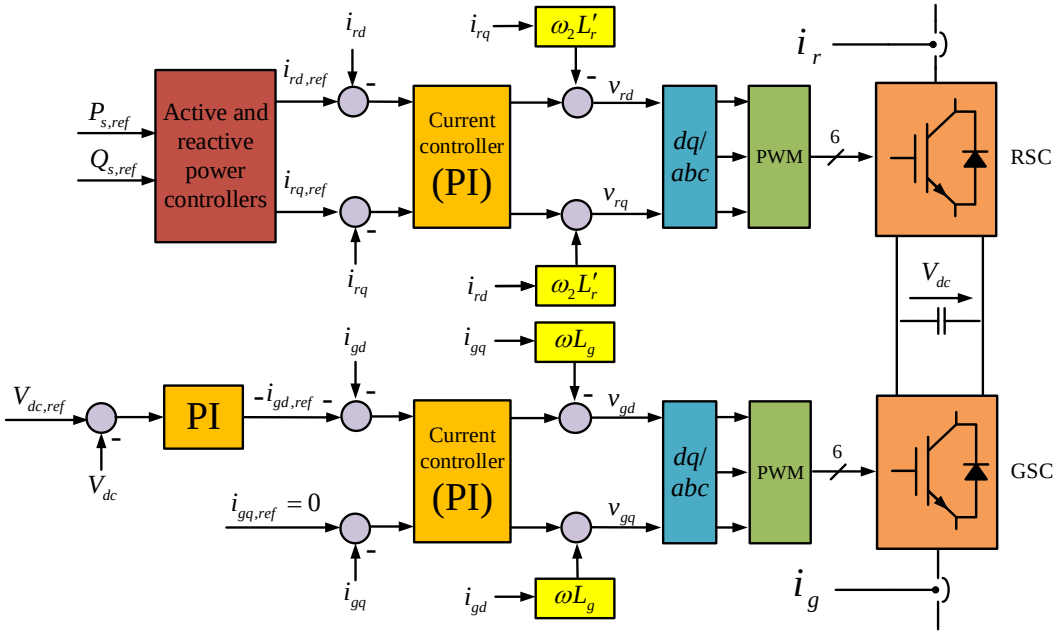


Fig. 2: Schematic diagram of PI controllers with PWM structure for RSC and GSC.

for an appropriate selected switching frequency is kept low enough by means of a filter. This filter which is located in the grid side is a series RL low pass filter which consists of R_g and L_g. The state space representations of grid side filter are obtained as

$$\begin{cases} \dfrac{di_{gd}}{dt} = -\dfrac{\omega_b R_g}{L_g} i_{gd} + \omega_b \omega i_{gq} - \dfrac{\omega_b}{L_g} v_{sd} + \dfrac{\omega_b}{L_g} v_{gd} \\ \dfrac{di_{gq}}{dt} = -\dfrac{\omega_b R_g}{L_g} i_{gq} - \omega_b \omega i_{gd} - \dfrac{\omega_b}{L_g} v_{sq} + \dfrac{\omega_b}{L_g} v_{gq} \end{cases} \quad (11)$$

Considering synchronous reference frame with the d axis aligned with the stator voltage space vector, the dynamic model of DC-link is described by the following instantaneous power balance.

$$CV_{dc} \frac{dV_{dc}}{dt} = -P_r - v_{sd} i_{gd} - P_{loss} \quad (12)$$

Fig. 2 shows the conventional control system of RSC and GSC. In the RSC, the d and q components of rotor current reference value are generated using active and reactive power controllers. The difference between these reference components and measured values of rotor current are applied to the current controller. v_{rd} and v_{rq} are obtained from the current controller and converted to abc quantities. Then, IGBT gate drive signals are generated using PWM switching technique.

The control system in GSC keeps the DC-link voltage constant as its main purpose. Taking into consideration the vector control oriented with stator voltage, $i_{gd,ref}$ controls the DC-link voltage. Reactive power which is injected into the power grid by GSC can be stated as

$$Q_g = Im\left\{v_{sdq} i_{gdq}^*\right\} = v_{sq} i_{gd} - v_{sd} i_{gq} \quad (13)$$

Accordingly, the reactive power is controlled by $i_{gq,ref}$ and can be injected into the grid. In this study, we set $Q_{g,ref} = 0$.

Review of SMC for Fault Ride Through

SMC is a nonlinear control technique with several advantages such as simplicity, robustness against system uncertainties and disturbances originated from external source, as well as good dynamical response. The SMC structure consists of equivalent control vector obtained in regard to the system mathematical model and switching part of control vector. In the SMC method, sliding surfaces are defined and the system states are pushed to their desired values. In [28], the sliding surfaces for RSC control have been considered the error between the measured and reference rotor currents as follows:

$$\begin{cases} s_{rd} = i_{rd,ref} - i_{rd} \\ s_{rq} = i_{rq,ref} - i_{rq} \end{cases} \quad (14)$$

The equivalent rotor voltages are obtained by supposing the sliding surface derivatives to be zero.

$$\begin{cases} \dfrac{ds_{rd}}{dt} = \dfrac{di_{rd,ref}}{dt} - \dfrac{di_{rd}}{dt} \\ \dfrac{ds_{rq}}{dt} = \dfrac{di_{rq,ref}}{dt} - \dfrac{di_{rq}}{dt} \end{cases} \quad (15)$$

The equivalent values of rotor voltages obtained by substituting (6) into (15) are stated as follows.

$$\begin{cases} v_{rd}^{eq} = \dfrac{L_r'}{\omega_b}\dfrac{di_{rd,ref}}{dt} + \left(R_r' i_{rd} - \omega_2 L_r' i_{rq} + E_d\right) \\ v_{rq}^{eq} = \dfrac{L_r'}{\omega_b}\dfrac{di_{rq,ref}}{dt} + \left(R_r' i_{rq} + \omega_2 L_r' i_{rd} + E_q\right) \end{cases} \quad (16)$$

The switching rotor voltages are properly designed so that the derivative of Lyapunov function, which is considered as $V = \frac{1}{2}s_i^2$, $i = rd, rq$, becomes negative-definite as follows:

$$\begin{cases} v_{rd}^{s} = k_{rd}\, sign\left(s_{rd}\right) & , k_{rd} \text{ is positive} \\ v_{rq}^{s} = k_{rq}\, sign\left(s_{rq}\right) & , k_{rq} \text{ is positive} \end{cases} \quad (17)$$

The rotor voltage consists of two parts: the equivalent value and the switching value given as

$$\begin{cases} v_{rd} = v_{rd}^{eq} + v_{rd}^{s} \\ v_{rq} = v_{rq}^{eq} + v_{rq}^{s} \end{cases} \quad (18)$$

Fig. 3 depicts the schematic diagram of the SMC used in the RSC control structure which has been proposed in [28]. It is adjusted according to equations in per unit values that have been presented in this paper.

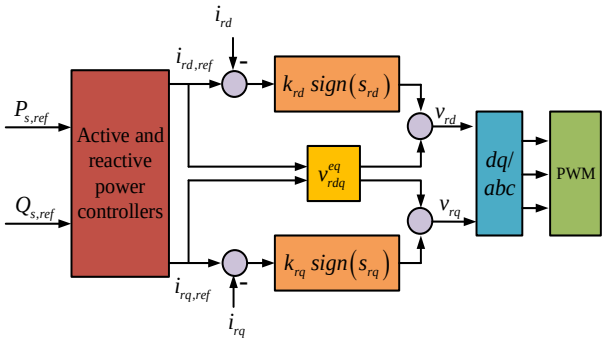


Fig. 3: Schematic diagram of the SMC used in RSC control structure.

In the GSC control structure, SMC sets the dq components of the GSC current by considering state space representations of the grid side filter, (11), and thus, defining sliding surfaces as follows:

$$\begin{cases} s_{gd} = i_{gd,ref} - i_{gd} \\ s_{gq} = i_{gq,ref} - i_{gq} \end{cases} \quad (19)$$

$$\begin{cases} \dfrac{ds_{gd}}{dt} = \dfrac{di_{gd,ref}}{dt} - \dfrac{di_{gd}}{dt} \\ \dfrac{ds_{gq}}{dt} = \dfrac{di_{gq,ref}}{dt} - \dfrac{di_{gq}}{dt} \end{cases} \quad (20)$$

$$\begin{cases} v_{gd}^{eq} = \dfrac{L_g}{\omega_b}\dfrac{di_{gd,ref}}{dt} + \left(R_g i_{gd} - \omega L_g i_{gq} + v_{sd}\right) \\ v_{rq}^{eq} = \dfrac{L_g}{\omega_b}\dfrac{di_{gq,ref}}{dt} + \left(R_g i_{gq} + \omega L_g i_{gd} + v_{sq}\right) \end{cases} \quad (21)$$

$$\begin{cases} v_{gd}^{s} = k_{gd}\, sign\left(s_{gd}\right) & , k_{gd} \text{ is positive} \\ v_{gq}^{s} = k_{gq}\, sign\left(s_{rq}\right) & , k_{gq} \text{ is positive} \end{cases} \quad (22)$$

$$\begin{cases} v_{gd} = v_{gd}^{eq} + v_{gd}^{s} \\ v_{gq} = v_{gq}^{eq} + v_{gq}^{s} \end{cases} \quad (23)$$

The block diagram of the SMC applied to the GSC, which has been proposed in [28], has been adjusted based on equations in per unit values within this paper (see Fig. 4).

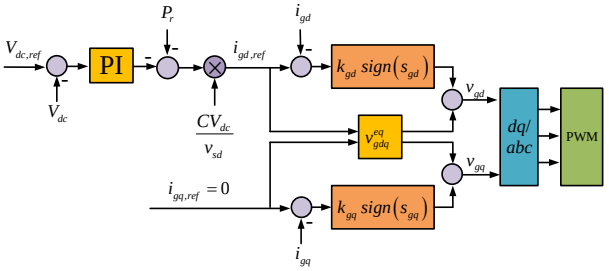


Fig. 4: Schematic diagram of the SMC used in GSC control structure.

The Proposed Predictive Approach

The MPC schemes for electromagnetic torque and current control in RSC and DC-link voltage and current control in GSC, which are on the basis of system models for predicting the action of state variables at the next sampling time and generating switching drive signals of power converters, are presented in this section. Firstly, the predicted value of a state variable is computed for all switching states which can exist. Then, the appropriate switching state which minimizes a cost function is chosen. Considering the discrete nature of power converters, the finite number of switching states, and the fast microprocessors, online minimization of the cost function is possible [31], [32]. It is essential to note that stability analysis of MPC controlled power converters has been presented in [33], [34] by considering the cost function of MPC as a candidate Lyapunov function. Fig. 5 illustrates the schematic diagram of the proposed MPC-based controllers' design for RSC and GSC.

A. MPC-Based Control of RSC

Fig. 6 shows an RSC that is assumed to be a three-phase converter with two power switches for each phase. At any specific time, only one switch is permitted to operate. The switching signals S_a , S_b , and S_c can be defined as

$$S_a = \begin{cases} 1 & if \;\; S_1 \;\; on \;\; and \;\; S_4 \;\; off \\ 0 & if \;\; S_1 \;\; off \;\; and \;\; S_4 \;\; on \end{cases} \quad (24)$$

$$S_b = \begin{cases} 1 & if \;\; S_2 \;\; on \;\; and \;\; S_5 \;\; off \\ 0 & if \;\; S_2 \;\; off \;\; and \;\; S_5 \;\; on \end{cases} \quad (25)$$

$$S_c = \begin{cases} 1 & if \;\; S_3 \;\; on \;\; and \;\; S_6 \;\; off \\ 0 & if \;\; S_3 \;\; off \;\; and \;\; S_6 \;\; on \end{cases} \quad (26)$$

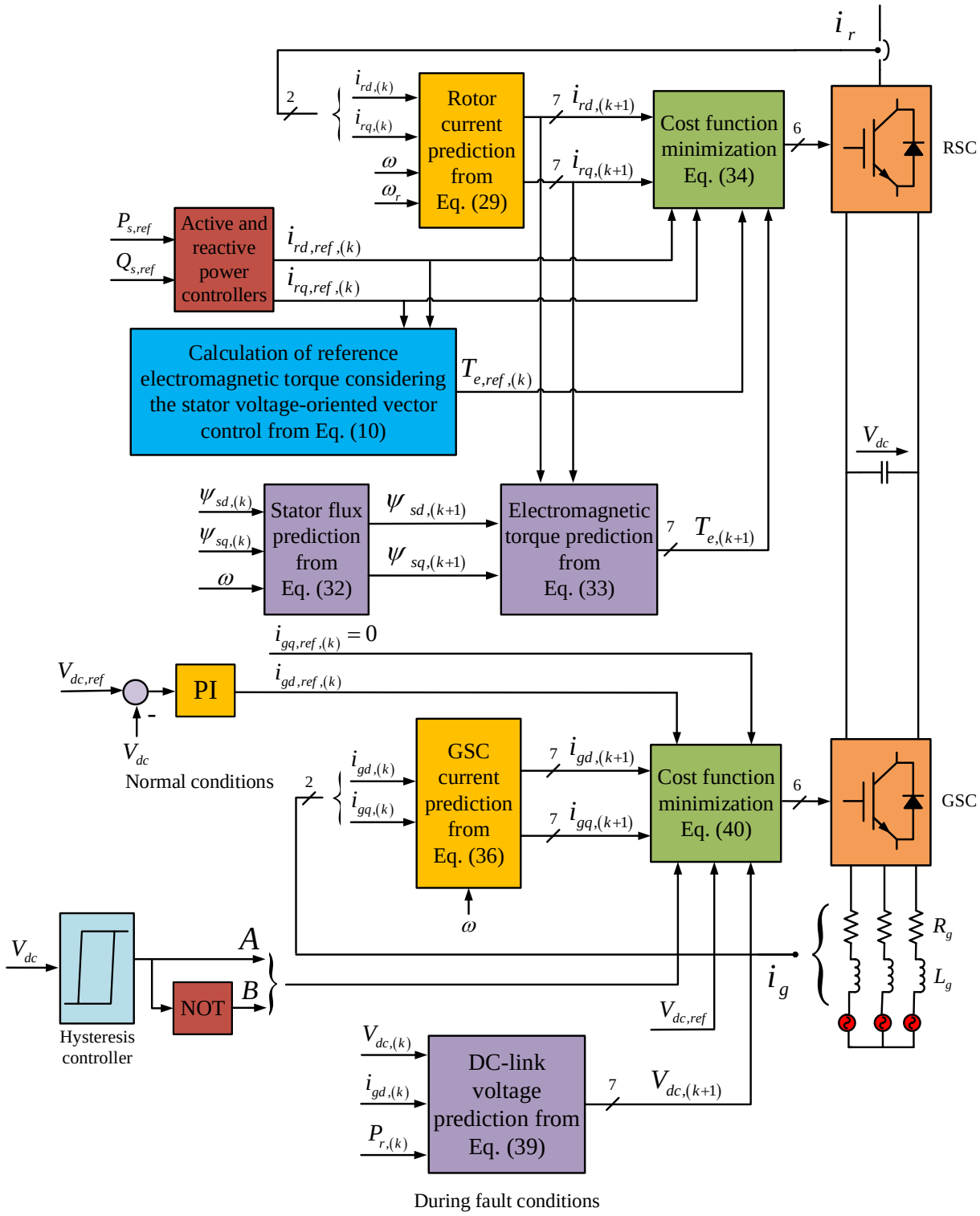


Fig. 5: Schematic diagram of the improved MPC implemented in RSC and GSC control structures.

The switching states of the three-phase two-level converter shown in Fig. 6 are given in Table 1. The converter's output voltage vector is described as follows:

$$\boldsymbol{v} = \frac{2}{3}\left(v_{aN} + \boldsymbol{a}v_{bN} + \boldsymbol{a}^2 v_{cN}\right) \quad (27)$$

where $\boldsymbol{a} = e^{j2\pi/3}$, and $v_{iN} = S_i V_{dc}$; $i = a,b,c$. Considering the possible combinations of switching signals, eight states and consequently seven different vectors of voltage are achieved (i.e. the switching states 1 and 8 in Table 1 present similar voltage vector which have been written in red.). In order to predict the action of dq components of rotor current, a discrete-time model in sampling time T_s is used which is obtained by a simple approximation of the derivatives. Forward Euler

approximation method is used to approximate the dq components of rotor current derivative as follows:

$$\begin{cases} \dfrac{di_{rd}}{dt} \approx \dfrac{i_{rd,(k+1)} - i_{rd,(k)}}{T_s} \\ \dfrac{di_{rq}}{dt} \approx \dfrac{i_{rq,(k+1)} - i_{rq,(k)}}{T_s} \end{cases} \quad (28)$$

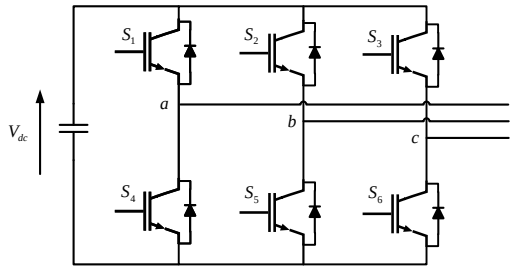


Fig. 6: Three-phase power converter circuit.

Table 1: Switching states of the three-phase two-level converter

	S_1	S_2	S_3	S_4	S_5	S_6
1	0	0	0	1	1	1
2	0	0	1	1	1	0
3	0	1	0	1	0	1
4	0	1	1	1	0	0
5	1	0	0	0	1	1
6	1	0	1	0	1	0
7	1	1	0	0	0	1
8	1	1	1	0	0	0

By replacement of (28) into (6), the dq components of rotor current at the next sampling time are given by

$$\begin{cases} i_{rd,(k+1)} = T_s\left(-\dfrac{\omega_b R'_r}{L'_r} i_{rd,(k)} + \omega_b \omega_2 i_{rq,(k)} - \dfrac{\omega_b}{L'_r} E_{d,(k)} + \dfrac{\omega_b}{L'_r} v_{rd,(k)}\right) + i_{rd,(k)} \\ i_{rq,(k+1)} = T_s\left(-\dfrac{\omega_b R'_r}{L'_r} i_{rq,(k)} - \omega_b \omega_2 i_{rd,(k)} - \dfrac{\omega_b}{L'_r} E_{q,(k)} + \dfrac{\omega_b}{L'_r} v_{rq,(k)}\right) + i_{rq,(k)} \end{cases} \quad (29)$$

The above equations are utilized for calculating a cost function which is given in the following paragraphs; it is defined so as to minimize the quadratic error between reference values of rotor current components and their predicted values. For simplicity, the reference rotor currents are considered not to change during each sampling time; thus:

$$i_{rd,ref,(k+1)} \approx i_{rd,ref,(k)}, \; i_{rq,ref,(k+1)} \approx i_{rq,ref,(k)} \quad (30)$$

Similar principle is used in predictive electromagnetic torque control. In other words, predictions are made for the future values of the stator flux and electromagnetic torque in this scheme. Forward Euler approximation is also considered to compute the dq components of the stator flux as follows:

$$\begin{cases} \dfrac{d\psi_{rd}}{dt} \approx \dfrac{\psi_{rd,(k+1)} - \psi_{rd,(k)}}{T_s} \\ \dfrac{d\psi_{rq}}{dt} \approx \dfrac{\psi_{rq,(k+1)} - \psi_{rq,(k)}}{T_s} \end{cases} \quad (31)$$

The stator flux components at the next sampling time are calculated by substituting (31) into (5).

$$\begin{cases} \psi_{sd,(k+1)} = T_s \omega_b \left(v_{sd,(k)} - R_s i_{sd,(k)} + \omega \psi_{sq,(k)}\right) + \psi_{sd,(k)} \\ \psi_{sq,(k+1)} = T_s \omega_b \left(v_{sq,(k)} - R_s i_{sq,(k)} - \omega \psi_{sd,(k)}\right) + \psi_{sq,(k)} \end{cases} \quad (32)$$

Given the values of stator flux and rotor current components at the next sampling time and substituting them into (10), the electromagnetic torque prediction is obtained:

$$T_{e,(k+1)} = \frac{L_m}{L_s}\left(\psi_{sq,(k+1)} i_{rd,(k+1)} - \psi_{sd,(k+1)} i_{rq,(k+1)}\right) \quad (33)$$

Similar to reference rotor currents, the reference electromagnetic torque at the next sampling time is assumed to be equal with the present sampling time.

The switching state is selected corresponding to minimum cost function for the next sampling time, and hence applied to RSC in order to achieve an appropriate electromagnetic torque and rotor current regulation. The cost function is defined as:

$$g_1 = \alpha\left[\left(i_{rd,ref,(k)} - i_{rd,(k+1)}\right)^2 + \left(i_{rq,ref,(k)} - i_{rq,(k+1)}\right)^2\right] + \beta\left[\left(T_{e,ref,(k)} - T_{e,(k+1)}\right)^2\right] \quad (34)$$

where, α and β are weighting factors and in this paper have been assigned 0.3 and 0.7, respectively by trial and error to improve both transient rotor current and transient electromagnetic torque during fault conditions. Due to the dependence of the electromagnetic torque on the rotor current, the rotor current can be indirectly controlled by controlling the electromagnetic torque. Hence, the weighting factor of electromagnetic torque term is considered to be larger than the weighting factor of the rotor current term in the (34). Based on [32], weighting factors are design parameters and adjusting of these factors depends on terms of the cost function. In other words, each term in the cost function is multiplied by a weighting factor to allow balancing of the different units and magnitudes of the controlled variables and to control their relative importance. A systematic way to determine these parameters is still a challenge and an open topic for research.

B. MPC-Based Control of GSC

Similar to RSC topology, GSC is considered to be a three-phase two-level converter with seven different vectors of voltage (Fig. 6). GSC current action is predicted by following approximation equations regarding to current derivative dq components:

$$\begin{cases} \dfrac{di_{gd}}{dt} \approx \dfrac{i_{gd,(k+1)} - i_{gd,(k)}}{T_s} \\ \dfrac{di_{gq}}{dt} \approx \dfrac{i_{gq,(k+1)} - i_{gq,(k)}}{T_s} \end{cases} \quad (35)$$

The predicted components of GSC current are obtained by substituting (35) into (11) as

$$\begin{cases} i_{gd,(k+1)} = T_s\left(-\dfrac{\omega_b R_g}{L_g} i_{gd,(k)} + \omega_b \omega i_{gq,(k)} - \dfrac{\omega_b}{L_g} v_{sd,(k)} + \dfrac{\omega_b}{L_g} v_{gd,(k)}\right) + i_{gd,(k)} \\ i_{gq,(k+1)} = T_s\left(-\dfrac{\omega_b R_g}{L_g} i_{gq,(k)} - \omega_b \omega i_{gd,(k)} - \dfrac{\omega_b}{L_g} v_{sq,(k)} + \dfrac{\omega_b}{L_g} v_{gq,(k)}\right) + i_{gq,(k)} \end{cases} \quad (36)$$

Since DC-link dynamics is nonlinear, the conventional PI control for DC-link voltage regulation will fail to operate properly considering uncertainties and voltage dips. Hence, during fault conditions, predictive DC-link voltage control is designed based on DC-link dynamics equation (12) regardless of GSC switching losses as follows:

$$CV_{dc}\frac{dV_{dc}}{dt} = -P_r - v_{gd} i_{gd} \quad (37)$$

Based on forward Euler approximation, the derivative of DC-link voltage can be written as

$$\frac{dV_{dc}}{dt} \approx \frac{V_{dc,(k+1)} - V_{dc,(k)}}{T_s} \quad (38)$$

The predicted DC-link voltage is calculated by substituting (38) into (37).

$$V_{dc,(k+1)} = T_s\left(-\frac{P_{r,(k)}}{CV_{dc}} - \frac{v_{gd,(k)} i_{gd,(k)}}{CV_{dc}}\right) + V_{dc,(k)} \quad (39)$$

Similar to MPC scheme in RSC, changes of the reference values in one sampling time are ignored.

The switching state of GSC is obtained on the basis of the defined cost function (40) in order to return the minimum value for the next sample time:

$$g_2 = A\left(V_{dc,ref} - V_{dc,(k+1)}\right)^2 + B\left(i_{gd,ref,(k)} - i_{gd,(k+1)}\right)^2 + \left(i_{gq,ref,(k)} - i_{gq,(k+1)}\right)^2 \quad (40)$$

As shown in Fig. 5, A and B are obtained from a hysteresis controller which its output takes zero or one values. It should be noted that according to the nominal value of V_{dc} (i.e. 1150 V) the upper and lower bands of the hysteresis controller have been considered 1165 V and 1155 V, respectively. When the DC-link voltage exceeds the upper band, the output value of the hysteresis controller will be one. If the DC-link voltage is less than the lower limit, the output value of the hysteresis controller is zero. In fact, at normal conditions the cost function (40) is the quadratic error between reference values of GSC current components and their predicted values. Whereas, during fault conditions, this cost function becomes the summation of the quadratic error between q component of predicted GSC current and its reference, and the quadratic error between predicted DC-link voltage and its reference.

Results and Discussion

The test system shown in Fig. 1 is modeled and simulated in the MATLAB-SIMULINK environment to investigate the performance of the proposed approach in FRT capability of the grid-connected DFIG-based WT with regard to the U.S. grid code stated by the Federal Energy Regulatory Commission, FERC. Fig. 7 shows the voltage dip ride through curve as specified by the FERC grid code. Accordingly, WTs must provide FRT support and remain connected to power grid in fault conditions with 85% depth and 600 ms duration in the point of common coupling (PCC) voltage [6]. In this study, the proposed control objectives are to restrict the current of rotor winding and the DC-link voltage to the ranges of 2 pu and 1.2 times of DC-link rated voltage value, respectively [35].

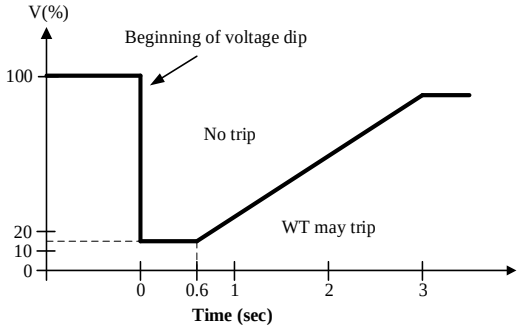


Fig. 7: Low voltage ride through (LVRT) requirement presented by FERC.

Parameters of the power grid, synchronous generator and DFIG-based WT shown in Fig. 1 are given in Tables 2-4. At time t=1 s, a 600 ms three-phase short-circuit fault has been simulated in the transmission line, which drops the PCC voltage to 15% of its rated value. Fig. 8 shows the transient response of DFIG-based WT during the fault.

Simulations are performed for a constant wind speed equal to 12 m/s. Results are compared with PI control and SMC methods which utilize PWM to implement the desired control.

For proper impartial comparison between conventional PWM and switching based on MPC theory, the same average of switching frequency has been considered in the latter one. Rotor current and DC-link voltage intense transients can be seen as the fault occurs and as it is cleared. When the conventional PI method is used, at the fault occurrence moment, the rotor current reaches 3.5 pu as a consequence of the magnetic coupling between stator and rotor.

Table 2: Parameters of electrical power grid

Rated voltage				25 kV
Rated frequency				60 Hz
Parameters of transmission lines				
	Sequences			
	Positive			Zero
R(Ω / km)	0.1153			0.413
L(mH / km)	1.05			3.32
C(nF / km)	11.33			5.01
	# 1	# 2	# 3	# 4
Length(km)	20	15	15	30
Parameters of loads				
	P(kW)			Q(kVAr)
# 1	400			120
# 2	600			150

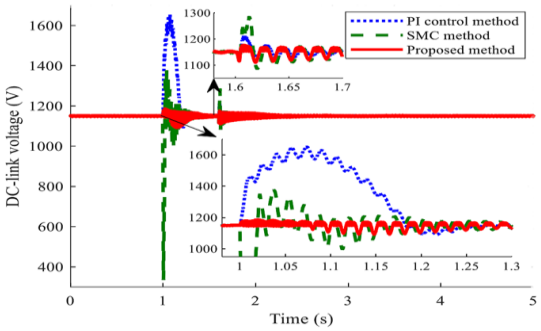
Table 3: Parameters of synchronous generator

Rated voltage	400 V
Rated power	85 kVA
Rated speed	1500 rpm
X_d	1.305 pu
X_d'	0.296 pu
X_d''	0.252 pu
X_q	0.474 pu
X_q''	0.243 pu
X_l	0.18 pu
T_d'	1.01 s
T_d''	0.053 s
T_{qo}''	0.1 s
Stator resistance	$2.8544 * 10^{-3}$ pu
Pole pairs	4
Inertia coefficient	3.2 s

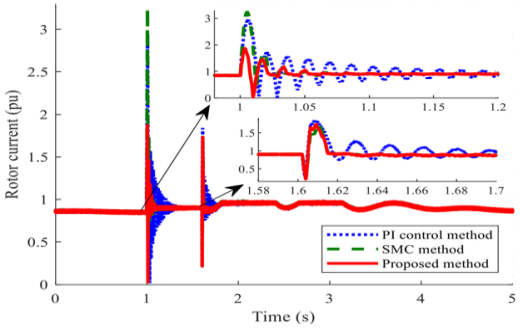
Table 4: Parameters of WT with DFIG

Rated power	1.5 MW
Rated value of v_s	575 V
Rated frequency	60 Hz
T_s	$5 * 10^{-6}$ s
Nominal wind speed	12 m/s
R_s	0.00706 pu
R_r	0.005 pu
Leakage inductance of stator	0.1716 pu
Leakage inductance of rotor	0.156 pu
L_m	2.9 pu
Pole pairs	3
Inertia constant	0.685 s
R_g	0.003 pu
Filter reactance in grid side	0.3 pu
Nominal value of V_{dc}	1150 V
C	10 mF

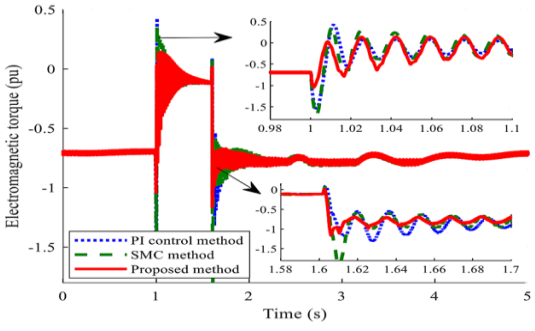
Furthermore, the sudden voltage drop prevents GSC from delivering the excess power to the power grid. Therefore, the excess power in RSC causes DC-link voltage fluctuations to exceed their permissible limit (around 1.4 pu in this case). Accordingly, there is the risk of damage to RSC and DC-link capacitor. SMC method can limit the DC-link over-voltage bellow 1.2 times of DC-link rated voltage value, but it is incapable to restrict the rotor current within permissible limits. However, using the proposed method, peak oscillations of DC-link voltage and rotor current will be restricted to acceptable thresholds and ride through capability of DFIG-based WT will be ensured. Significant mechanical stresses are created due to electromagnetic torque fluctuations under low grid voltage conditions which consequently reduce the machine reliability. The performance of the proposed method in electromagnetic torque peak suppression can be observed in Fig. 8c. As shown in Figs. 8d and 8e, output active and reactive powers of DFIG-based WT by using the proposed method have better transient responses than PI control and SMC methods, especially at the fault clearance time.



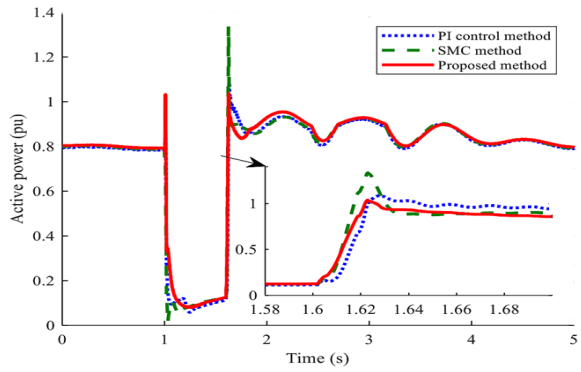
(a)



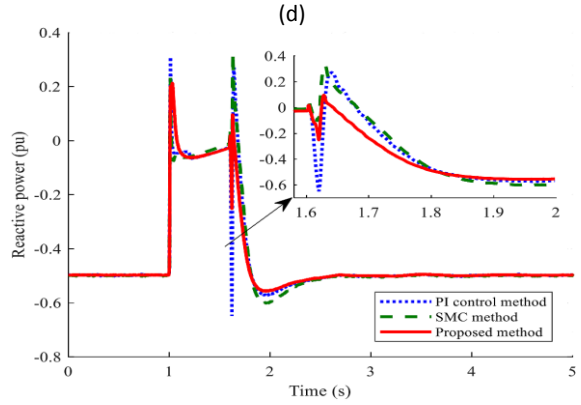
(b)



(c)

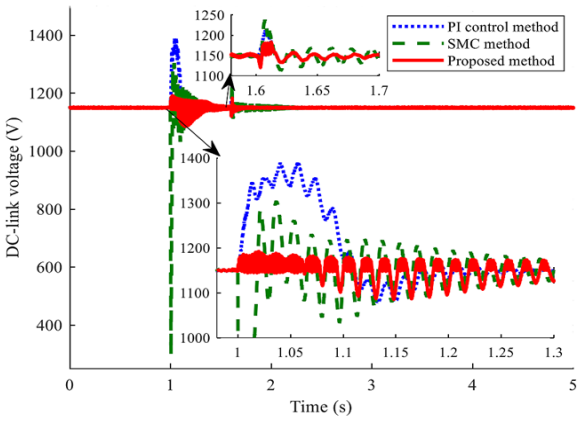


(d)

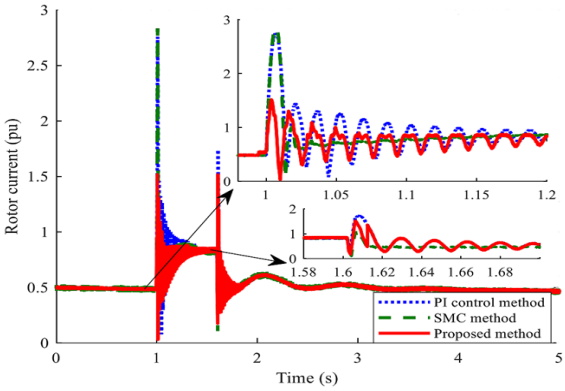


(e)

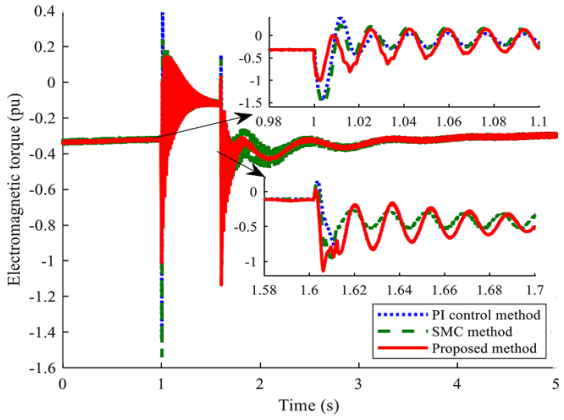
Fig. 8: Transient behavior of DFIG-based WT under wind speed of 12 m/s and an 85% three-phase fault: (a) DC-link voltage (V); (b) rotor current (pu); (c) electromagnetic torque (pu); (d) active power (pu); (e) reactive power (pu).



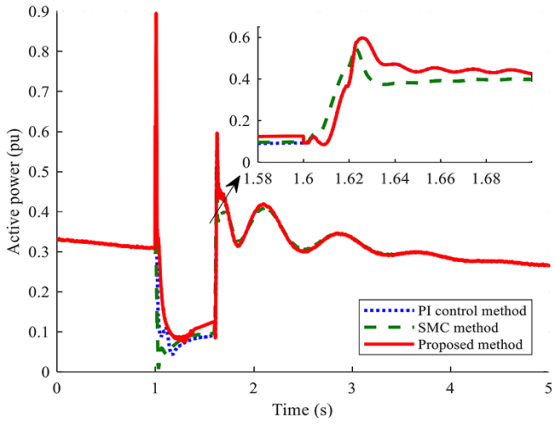
(a)



(b)



(c)



(d)

A. Operation with Variable Wind Speed

In order to prove the effectiveness of the proposed MPC-based control approach considering the variability of wind speed, three other case studies are conducted in this part.

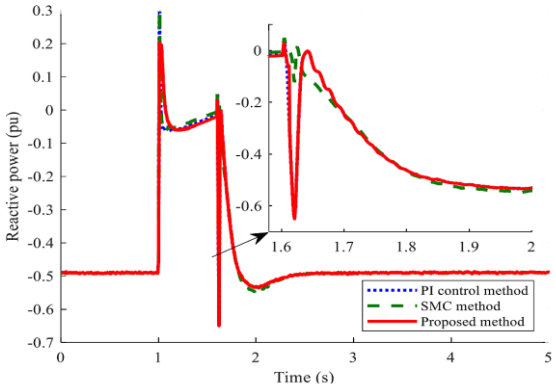
A three-phase fault occurs such that it leads to 85% voltage dip.

The simulated fault is cleared after 600 ms. Transient behavior of DFIG-based WT using wind speeds of 7 and 10 m/s are shown in Figs. 9 and 10, respectively.

Also, in order to investigate real wind turbulence in action, variable wind speed firstly starts with 12 m/s; then at time t=10 s a real turbulence term which is generated by Dryden velocity spectra [36] is applied.

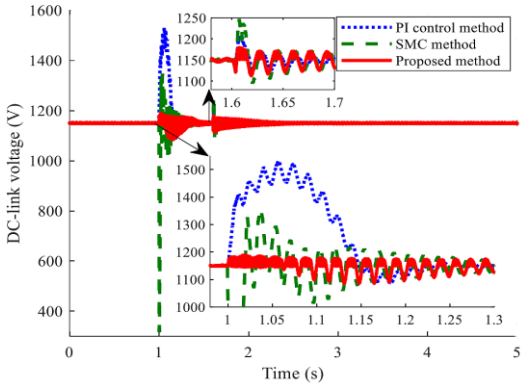
The wind speed variations and its effects on transient behavior of DFIG-based WT are illustrated in Figs. 11 and 12, respectively. It is observed from Figs. 9, 10, and 12 that the SMC method cannot limit the rotor current to 2 pu, but it keeps the DC-link voltage within the permissible range.

Also, the proposed FRT enhancement approach is capable to constrain the DC-link voltage, rotor current and electromagnetic torque within specified bounds, and thus, the FRT capability of DFIG-based WT is improved.

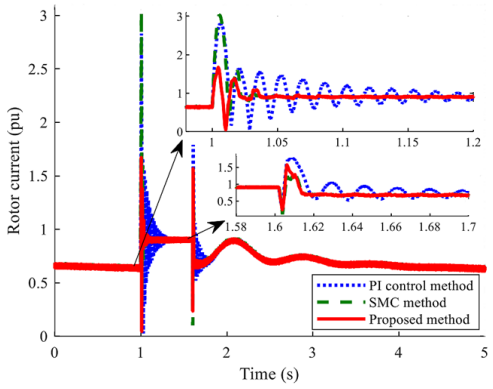


(e)

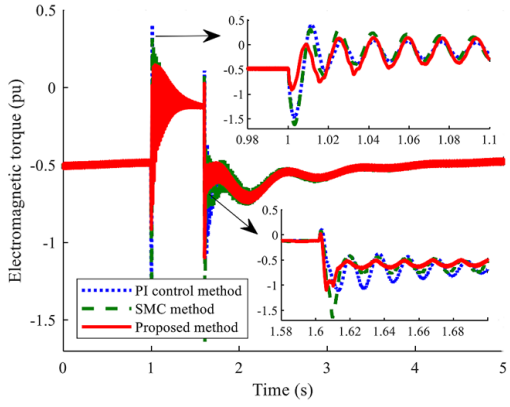
Fig. 9: Transient behavior of DFIG-based WT under wind speed of 7 m/s and an 85% three-phase fault: (a) DC-link voltage (V); (b) rotor current (pu); (c) electromagnetic torque (pu); (d) active power (pu); (e) reactive power (pu).



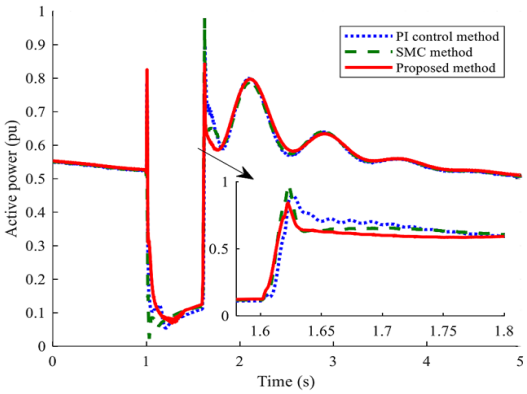
(a)



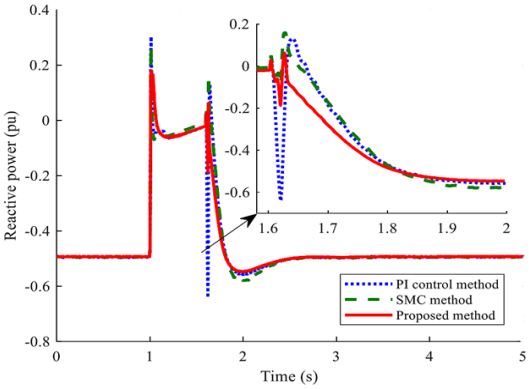
(b)



(c)



(d)



(e)

Fig. 10: Transient behavior of DFIG-based WT under wind speed of 10 m/s and an 85% three-phase fault: (a) DC-link voltage (V); (b) rotor current (pu); (c) electromagnetic torque (pu); (d) active power (pu); (e) reactive power (pu).

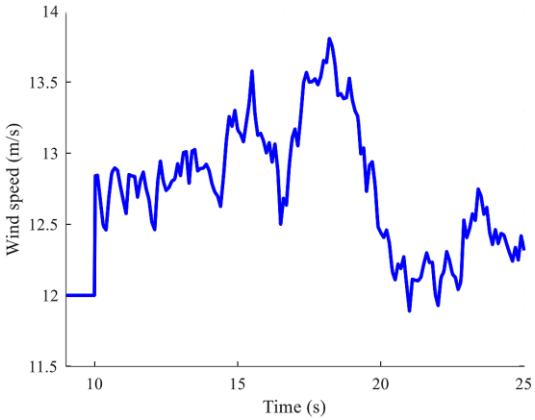
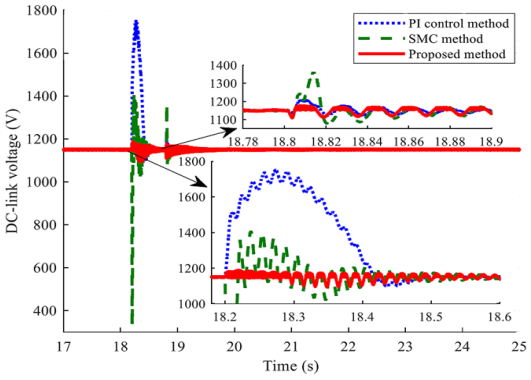


Fig. 11: Wind profile (m/s).

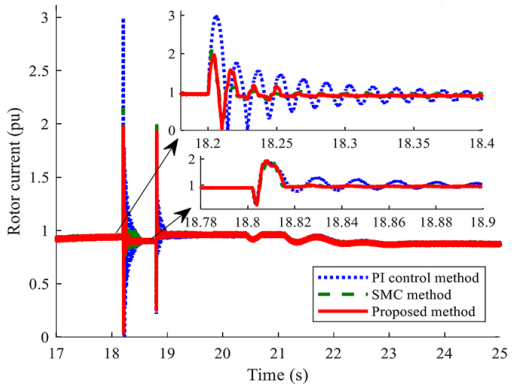
B. Operation with Single-Phase Fault

A short-circuit fault of single-phase to ground is tested to examine the robustness of the improved MPC schemes for enhancing FRT capability. Given the DFIG-based WT operating under wind speed of 12 m/s, the performance of coordinated MPC approach is compared to PI control and SMC methods which is shown in Fig. 13. The proposed method improves the peak DC-link voltage, rotor current, and electromagnetic torque by 5.53%, 35.33%, and 83.75%, respectively, and reduces

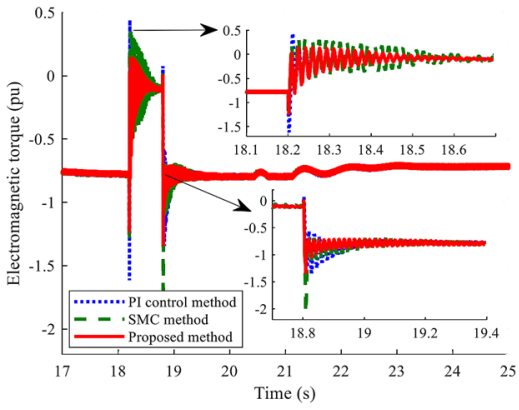
their corresponding oscillations in comparison with the results derived from PI and SMC applications for FRT capability enhancement.



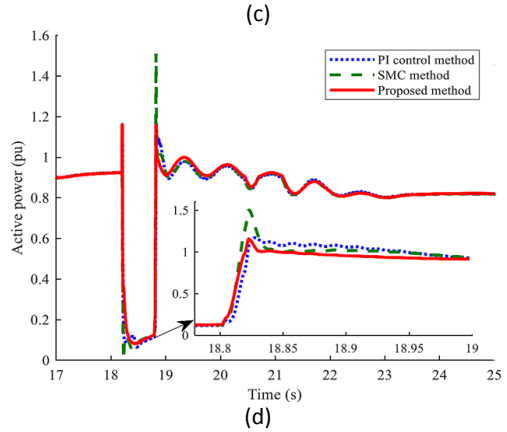
(a)



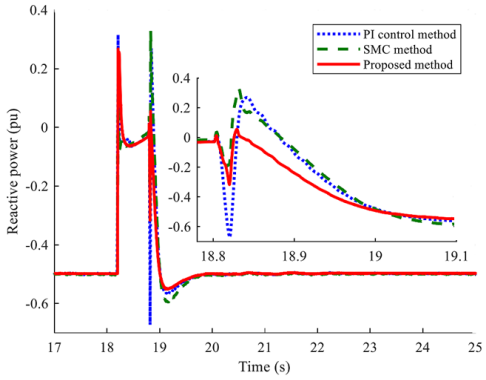
(b)



(c)

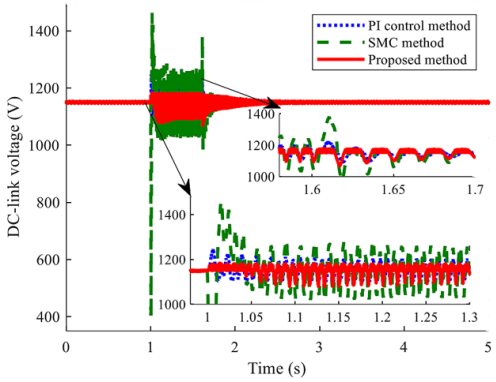


(d)

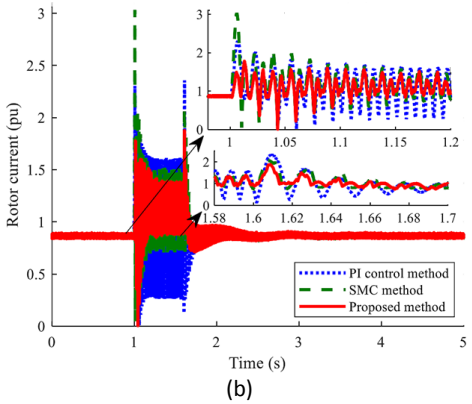


(e)

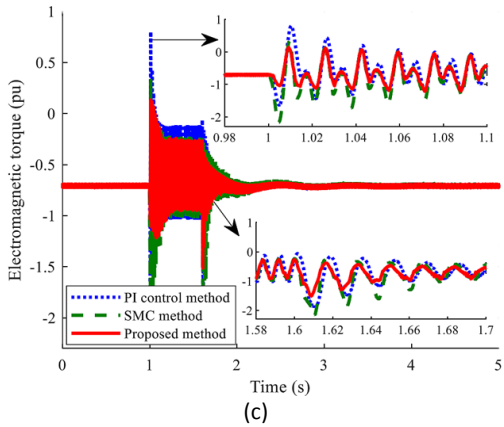
Fig. 12: Transient behavior of DFIG-based WT under variable wind speed and an 85% three-phase fault: (a) DC-link voltage (V); (b) rotor current (pu); (c) electromagnetic torque (pu); (d) active power (pu); (e) reactive power (pu).



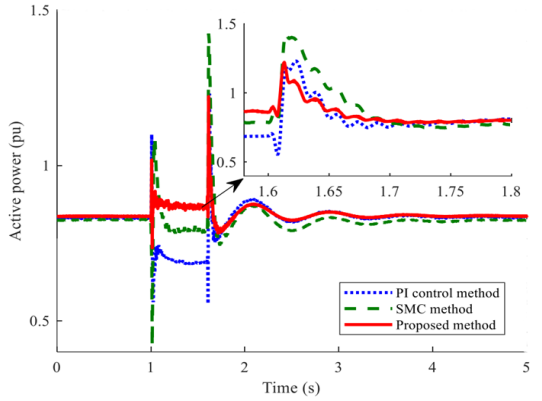
(a)



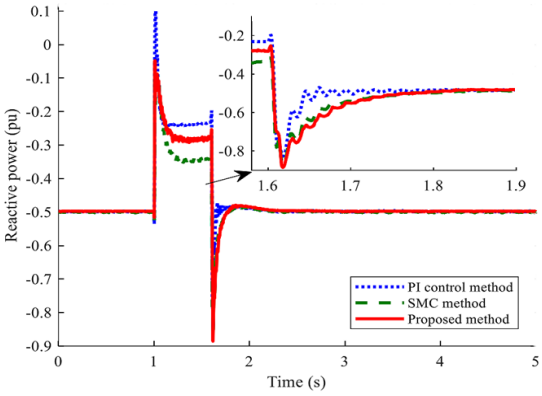
(b)



(c)



(d)

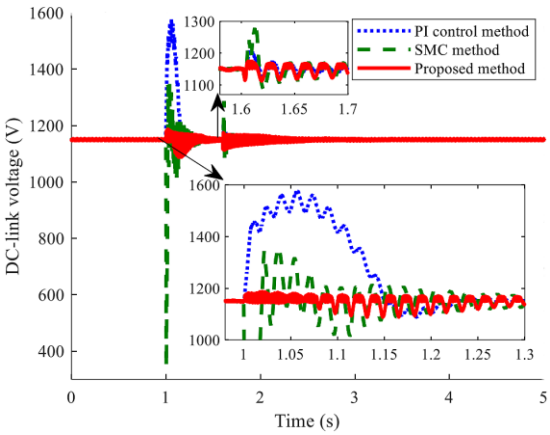


(e)

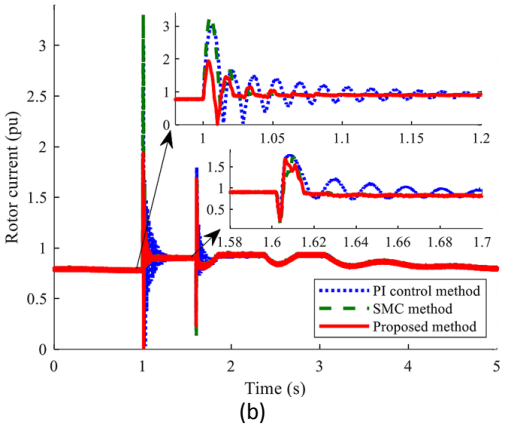
Fig. 13: Transient behavior of DFIG-based WT under wind speed of 12 m/s and single-phase fault: (a) DC-link voltage (V); (b) rotor current (pu); (c) electromagnetic torque (pu); (d) active power (pu); (e) reactive power (pu).

C. Operation with Variation in DFIG Parameters

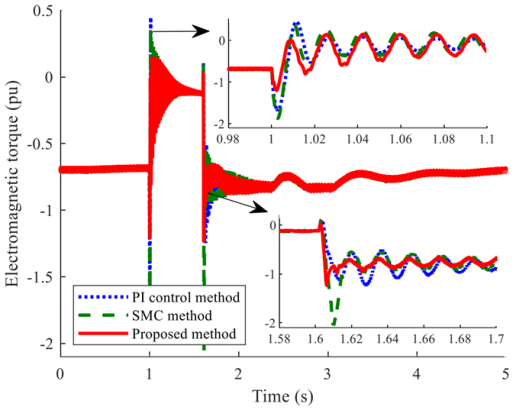
In this part, the simulation is performed when the magnetizing inductance, stator and rotor resistances are 1.5 and 0.5 times of their nominal values to investigate robust performance of the proposed control method against parameter uncertainties [21]. Figs. 14 and 15 depict the simulation results. In such cases, where DFIG parameters vary, the proposed MPC-based control scheme is more effective than both of the PI control and SMC methods in terms of FRT enhancement.



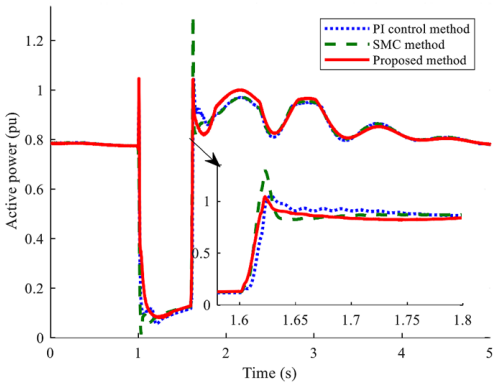
(a)



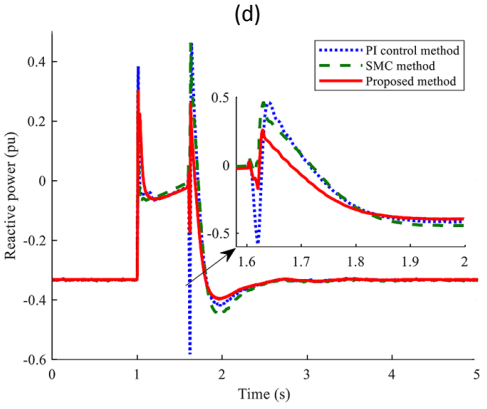
(b)



(c)

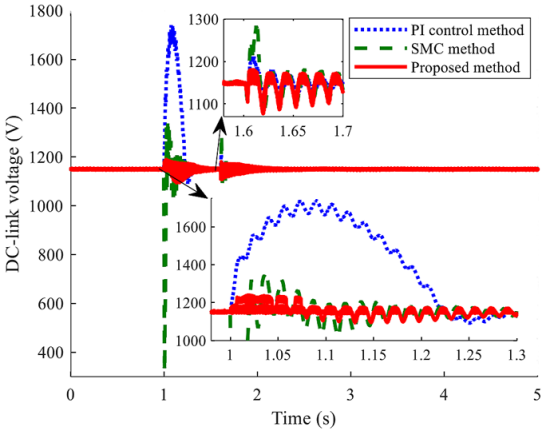


(d)

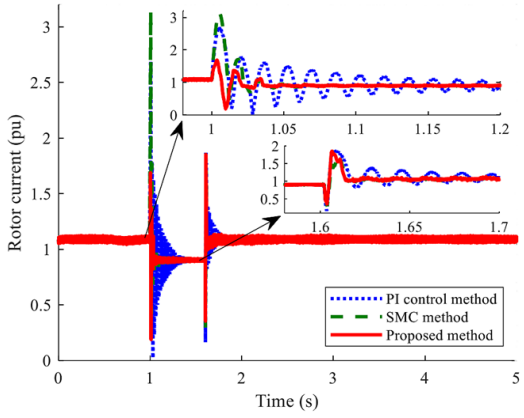


(e)

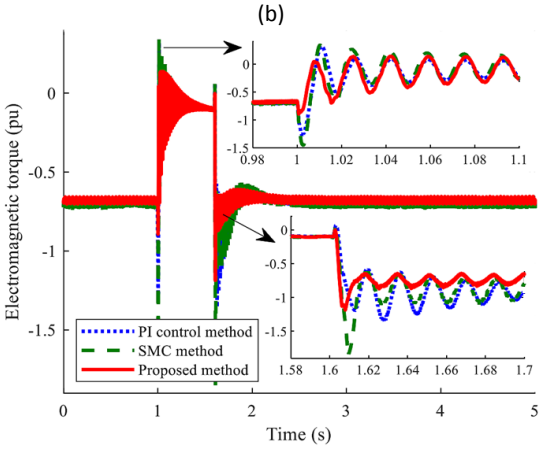
Fig. 14: Transient behavior of DFIG-based WT under wind speed of 12 m/s, an 85% three-phase fault and DFIG parameter variation: $R_s \rightarrow 1.5R_s$, $R_r \rightarrow 1.5R_r$, $L_m \rightarrow 1.5L_m$; (a) DC-link voltage (V); (b) rotor current (pu); (c) electromagnetic torque (pu); (d) active power (pu); (e) reactive power (pu).



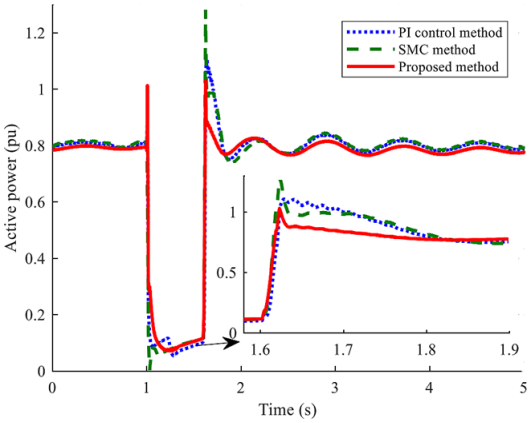
(a)



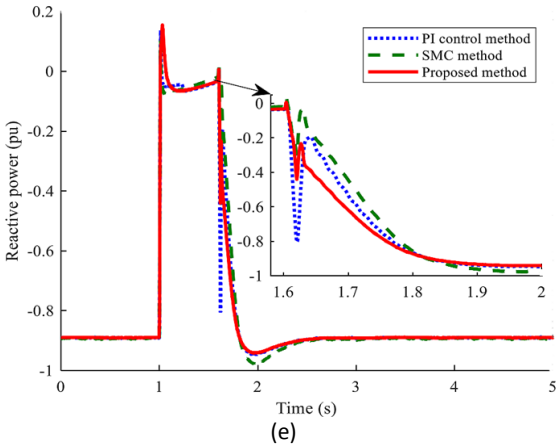
(b)



(c)



(d)

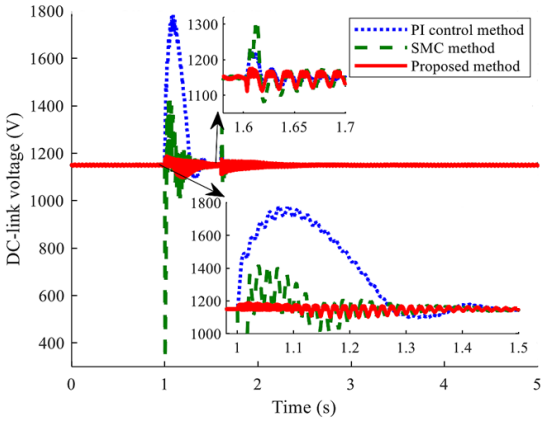


(e)

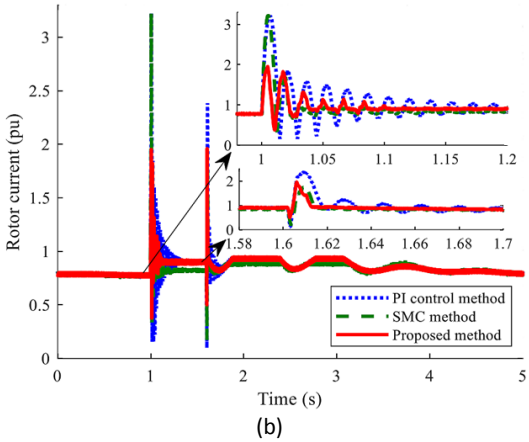
Fig. 15: Transient behavior of DFIG-based WT under wind speed of 12 m/s, an 85% three-phase fault and DFIG parameter variation: $R_s \rightarrow 0.5R_s$, $R_r \rightarrow 0.5R_r$, $L_m \rightarrow 0.5L_m$; (a) DC-link voltage (V); (b) rotor current (pu); (c) electromagnetic torque (pu); (d) active power (pu); (e) reactive power (pu).

D. Operation with a Severe Three-Phase Fault

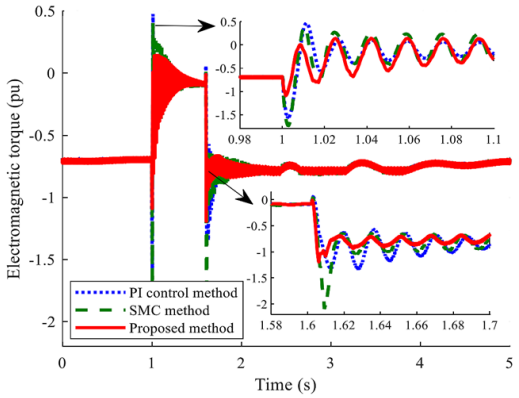
In order to study the DFIG transient behavior under a severe three-phase fault using the proposed method, a 90% three-phase dip in voltage is considered. In Fig. 16, simulation results show that using the coordinated MPC approach, the peak rotor current, electromagnetic torque and DC-link voltage are limited significantly.



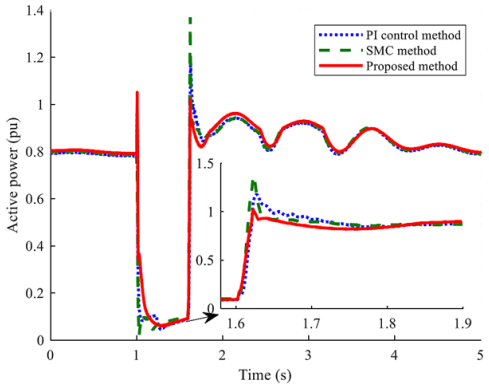
(a)



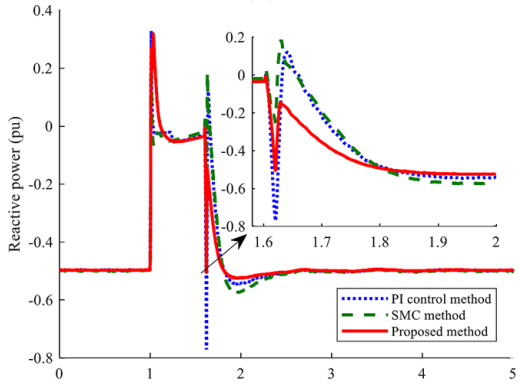
(b)



(c)



(d)



(e)

Fig. 16: Transient behavior of DFIG-based WT under wind speed of 12 m/s and a 90% three-phase fault: (a) DC-link voltage (V); (b) rotor current (pu); (c) electromagnetic torque (pu); (d) active power (pu); (e) reactive power (pu).

From the above different case studies provided in this section, it can be observed that, using the proposed method, the DC-link voltage and the rotor current are kept within acceptable limits. Moreover, it is worth noticing the fast dynamic behavior of the proposed method. This can be explained by the fact that MPC is a direct strategy that does not require an inner PI control loop for modulators. Hence, there is no bandwidth limitation for the electromagnetic torque dynamics. Although the proposed control scheme requires lots of calculations compared to conventional methods, fortunately, the today performance of modern microprocessors is sufficiently high to make this approach practical.

Conclusion

This paper proposed a novel MPC-based control strategy to improve the FRT capability of grid-connected DFIG-based WTs. Simulation results illustrated that the proposed control scheme is able to effectively reduce the peak values of DC-link voltage, rotor current and electromagnetic torque, while maintaining their values within the acceptable threshold. As a consequence of the inherent fast dynamics of the proposed control method, it also reduces system oscillations during fault conditions. Accordingly, the DFIG-based WT successfully rides through grid faults and provides continuous active/reactive power for the grid during and post faults without requiring any auxiliary hardware protection devices, ensuring the compliance to grid code requirements.

Finally, the robustness, effectiveness, and proper operation of the proposed control strategy over PI control and SMC techniques with PWM switching algorithm was demonstrated by conducting several different case study simulations including: variable wind speeds, severe voltage dips, single-phase faults, and DFIG parameters variations. It should be noted that maximum peak values of DC-link voltage by using PI control, SMC and the proposed methods in different case studies were 1783, 1463 and 1190 V, respectively. The PI control, SMC and the proposed methods provided 3.23, 3.3 and 1.95 pu values, respectively as maximum peak of rotor current. Also, the maximum peak values of electromagnetic torque by using PI control, SMC and the proposed MPC strategies were 0.8, 0.4 and 0.14 pu, respectively. To conclude, the proposed model predictive approach can be considered as a fast, robust and effective FRT crowbarless solution for grid-connected DFIG-based WTs.

Author Contributions

Z. Dehghani Arani, Prof. S. A. Taher, and M. H. Karimi proposed and designed the improved predictive control approach for DFIG-based wind turbine. Z. Dehghani Arani, and M. H. Karimi collected the data and wrote the original draft of the manuscript. Prof. S. A. Taher, and Z. Dehghani Arani carried out the data analysis. Prof. S. A. Taher, and Dr. M. Rahimi interpreted the results as well as reviewed and edited the manuscript.

Acknowledgment

The authors gratefully acknowledge the vice-chancellor for research and technology of University of Kashan for the supports.

Conflict of Interest

The authors declare that there is no conflict of

interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancy have been completely observed by the authors.

Abbreviations

α , β	Weighting factors
ψ	Flux
ω	dq reference frame's speed
ω_b	Base value of angular frequency (in this study, we considered $\omega = \omega_b$)
ω_r	Rotor speed
$\omega_2 = \omega - \omega_r$	Rotor slip frequency
A , B	Zero or one values obtained from the hysteresis controller
ANFIS	Adaptive neuro-fuzzy inference system
C	DC-link capacitance
DFIG	Doubly fed induction generator
DVR	Dynamic voltage restorer
e	Error between the measured and reference values of parameters
E	Rotor back-EMF voltage
FRT	Fault ride through
g_1, g_2	Cost functions in RSC and GSC control structures respectively
GSC	Grid side converter
i	Current
k	SMCs' parameters
L_g	Filter inductance used in grid side
L_m	Magnetizing inductance
L_r	Self-inductance of rotor
L'_r	Transient inductance
L_s	Self-inductance of stator
LVRT	Low voltage ride through
MPC	Model predictive control
P	Active power
PCC	Point of common coupling
PI	Proportional-plus-integral
P_{loss}	Total conducting and switching losses of the GSC
P_r	Rotor instantaneous input power
PWM	Pulse-width modulation
Q	Reactive power
R_g	Filter resistance used in grid side
R_r	Resistance of rotor
R'_r	Transient resistance
R_s	Resistance of stator
RSC	Rotor side converter
s	Sliding surfaces
S_a , S_b , S_c	Switching signals
SC	Superconducting coil
SFCL	Superconducting fault current limiter
SMC	Sliding mode control
SMES-FCL	Superconducting magnetic energy storage-fault current limiter
STATCOM	Static synchronous compensator
SVC	Static volt ampere reactive compensator
T_e	Electromagnetic torque
T_s	Sampling time
v	Voltage
V_{dc}	DC-link voltage
WECS	Wind energy conversion system
WT	Wind turbine

Subscripts

d, q	Synchronous dq reference frame
g	Quantities of grid side filter
(k)	Quantities at the k^{th} sampling time (present sampling time)
$(k+1)$	Quantities at the $k+1^{th}$ sampling time (next sampling time)
$nominal$	Nominal quantity
r	Quantities of rotor
ref	Reference quantities
s	Quantities of stator

Superscripts

eq	Equivalent control inputs
s	Switching control inputs

References

[1] E. J. N. Menezes, A. M. Araújo, N. S. B. da Silva, "A review on wind turbine control and its associated methods," J. Clean. Prod., 174: 945-953, 2018.

[2] M. J. Morshed, A. Fekih, "A new fault ride-through control for DFIG-based wind energy systems," Electr. Power Syst. Res., 146: 258-269, 2017.

[3] J. Morren, S. W. H. de Haan, "Ridethrough of wind turbines with doubly-fed induction generator during a voltage dip," IEEE Trans. Energy Convers., 20(2): 435-441, 2005.

[4] S. Wang, N. Chen, D. Yu, A. Foley, L. Zhu, K. Li, J. Yu, "Flexible fault ride through strategy for wind farm clusters in power systems with high wind penetration," Energy Convers. Manage., 93(3): 239-248, 2015.

[5] A. Jalilian, S. B. Naderi, M. Negnevitsky, M. Tarafdar Hagh, K. M. Muttaqi, "Low voltage ride-through enhancement of DFIG-based wind turbine using DC link switchable resistive type fault current limiter," Electr. Power Energy Syst., 86: 104-119, 2017.

[6] M. Tsili, S. Papathanassiou, "A review of grid code technical requirements for wind farms," IET Renew. Power Gen., 3(3): 308-332, 2009.

[7] D. Campos-Gaona, E. L. Moreno-Goytia, O. Anaya-Lara, "Fault ride-through improvement of DFIG-WT by integrating a two-degrees-of-freedom internal model control," IEEE Trans. Ind. Electron., 60(3): 1133-1145, 2013.

[8] W. Chen, D. Xu, N. Zhu, M. Chen, F. Blaabjerg, "Control of doubly fed induction generator to ride through recurring grid faults," IEEE Trans. Power Electron., 31(7): 4831-4846, 2016.

[9] X. Xiao, R. Yang, X. Chen, Z. Zheng, C. Li, "Enhancing fault ride-through capability of DFIG with modified SMES-FCL and RSC control," IET Gener. Transmiss. Distrib., 12(1): 258-266, 2018.

[10] S. I. Gkavanoudis, C. S. Demoulias, "Fault ride-through capability of a DFIG in isolated grids employing DVR and supercapacitor energy storage," Int. J. Electr. Power Energy Syst., 68: 356-363, 2015.

[11] Y. Kailasa Gounder, D. Nanjundappan, V. Boominathan, "Enhancement of transient stability of distribution system with SCIG and DFIG based wind farms using STATCOM," IET Renew. Power Gen., 10(8): 1171-1180, 2016.

[12] A. Safaei, B. Vahidi, S. H. Hosseinian, H. A. Abyaneh, "Fault ride-through capability improvement of doubly fed induction generator-based wind turbine using static volt ampere reactive compensator," AIP J. Renewable Sustainable Energy, 7(2), 2015.

[13] T. Karaipoom, I. Ngamroo, "Optimal superconducting coil integrated into DFIG wind turbine for fault ride through capability enhancement and output power fluctuation suppression," IEEE Trans. Sustain. Energy, 6(1): 28-42, 2015.

[14] S. Teimourzadeh, F. Aminifar, M. Davarpanah, J. M. Guerrero, "Macroprotections for microgrids: toward a new protection paradigm subsequent to distributed energy resource integration," IEEE Ind. Electron. Mag., 10(3): 6-18, 2016.

[15] S. B. Naderi, M. Negnevitsky, K. M. Muttaqi, "A modified DC chopper for limiting the fault current and controlling the DC-link voltage to enhance fault ride-through capability of doubly-fed induction-generator-based wind turbine," IEEE Trans. Ind. Appl., 55(2): 2021-2032, 2019.

[16] K. Du, X. Xiao, Y. Wang, Z. Zheng, C. Li, "Enhancing fault ride-through capability of DFIG-based wind turbines using inductive SFCL with coordinated control," IEEE Trans. Appl. Superconduct., 29(2): 1-6, 2019.

[17] Y. M. Alsmadi, L. Xu, F. Blaabjerg, A. J. P. Ortega, A. Y. Abdelaziz, A. Wang, Z. Albataineh, "Detailed investigation and performance improvement of the dynamic behavior of grid-connected DFIG-based wind turbines under LVRT conditions," IEEE Trans. Ind. Appl., 54(5): 4795-4812, 2018.

[18] M. J. Hossain, T. K. Saha, N. Mithulananthan, H. R. Pota, "Control strategies for augmenting LVRT capability of DFIGs in interconnected power system," IEEE Trans. Ind. Electron., 60(6): 2510-2522, 2013.

[19] D. Zhu, X. Zou, S. Zhou, W. Dong, Y. Kang, J. Hu, "Feedforward current references control for DFIG-based wind turbine to improve transient control performance during grid faults," IEEE Trans. Energy Convers., 33(2): 670-681, 2018.

[20] D. Xie, Z. Xu, L. Yang, J. Østergaard, Y. Xue, K. P. Wong, "A comprehensive LVRT control strategy for DFIG wind turbines with enhanced reactive power support," IEEE Trans. Power Syst., 28(3): 3302-3310, 2013.

[21] M. Rahimi, M. Parniani, "Transient performance improvement of wind turbines with doubly fed induction generators using nonlinear control strategy," IEEE Trans. Energy Convers., 25(2): 514-525, 2010.

[22] J. Liang, W. Qiao, R. G. Harley, "Feed-forward transient current control for low-voltage ride-through enhancement of DFIG wind turbines," IEEE Trans. Energy Convers., 25(3): 836-843, 2010.

[23] J. Liang, D. F. Howard, J. A. Restrepo, R. G. Harley, "Feed-forward transient compensation control for DFIG wind turbines during both balanced and unbalanced grid disturbances," IEEE Trans. Ind. Appl., 49(3): 1452-1463, 2013.

[24] A. J. Sguarezi Filho, E. R. Filho, "Model-based predictive control applied to the doubly-fed induction generator direct power control," IEEE Trans. Sustain. Energy, 3(3): 398-406, 2012.

[25] M. Soliman, O. P. Malik, D. T. Weswick, "Ensuring fault ride through for wind turbines with doubly fed induction generator: a model predictive control approach," in Proc. 18th IFAC World Conf.: pp. 1710-1715, 2011.

[26] M. Abdelrahem, M. H. Mobarak, R. Kennel, "Model predictive control for low-voltage ride through capability enhancement of DFIGs in variable-speed wind turbine systems," in Proc. of IEEE 9th Int. Conf. on Elect. and Computer Engineering,: 70-73, 2016.

[27] S. A. Taher, Z. Dehghani Arani, M. Rahimi, M. Shahidehpour, "Model predictive fuzzy control for enhancing FRT capability of DFIG-based WT in real-time simulation environment," Energy Syst., 9(4): 899-919, 2018.

[28] H. S. Naggar, A. S. Ahmed, M. M. Abd El-Aziz, "Low voltage ride through of doubly fed induction generator connected to the grid using sliding mode control strategy," Renew. Energy, 80: 583-594, 2015.

[29] S. A. Taher, Z. Dehghani Arani, M. Rahimi, M. Shahidehpour, "A new approach using combination of sliding mode control and feedback linearization for enhancing fault ride through capability of DFIG-based WT," Int. Trans. Electr. Energ. Syst., 28(10), 2018.

[30] M. Rahimi, "Coordinated control of rotor and grid sides converters in DFIG based wind turbines for providing optimal reactive power support and voltage regulation," Sustain. Energy Technol. Assess., 20: 47-57, 2017.

[31] S. Kouro, P. Cortes, R. Vargas, U. Ammann, J. Rodriguez, "Model predictive control-A simple and powerful method to control power converters," IEEE Trans. Ind. Electron., 56(6): 1826-1838, 2009.

[32] J. Rodriguez, P. Cortes, Predictive control of power converters and electrical drives, John Wiley & Sons, 2012.

[33] R. P. Aguilera, D. E. Quevedo, "On stability and performance of finite control set MPC for power converters," in Proc. Workshop Predictive Control Elect. Drives Power Electron.: 55-62, 2011.

[34] R. P. Aguilera, D. E. Quevedo, "Predictive control of power converters: Designs with guaranteed performance," IEEE Trans. Ind. Informat., 11(1): 53-63, 2015.

[35] A. H. Kasem, E. F. El-Saadany, H. H. El-Tamaly, M. A. A. Wahab, "An improved fault ride-through strategy for doubly fed induction generator-based wind turbines," IET Renew. Power Gener., 2(4): 201-214, 2008.

[36] C. K. Patel, H. T. Lee, I. M. Kroo, "Extracting energy from atmospheric turbulence," XXIX OSTIV Congress: 1-9, 2008.

Biographies

Zahra Dehghani Arani (M'17) was born in Kashan, Isfahan, Iran, in 1990. She received both her B.Sc. and M.Sc. degrees in electrical engineering from the University of Kashan, Kashan, Iran, in 2012 and 2015, respectively. She is currently pursuing Ph.D. degree of power systems engineering at University of Kashan, Isfahan, Iran. Her current research interests include renewable energy, micro/smart grids, fuzzy expert systems, predictive control, nonlinear control, multi-agent systems, and applications of artificial intelligence and parallel computing in power engineering.

Seyed Abbas Taher (SM' 17) was born in Kashan, Isfahan, Iran, in 1964. He received his B.Sc. degree in electrical engineering from the Amirkabir University of Technology, Tehran, Iran in 1988, and his M.Sc. and Ph.D. degrees in electrical engineering from the Tarbiat Modares University, Tehran, Iran, in 1991 and 1997, respectively. In 1996, he joined the faculty of engineering, University of Kashan, where he has been a Full Professor since 2016. His current research interests include power system optimization and control design, analysis of electrical machines, power quality, and renewable energy.

Mohammad Hossein Karimi was born in Faridan, Isfahan, Iran, in 1986. He was graduated with B.Sc. degree in power electric engineering from the Islamic Azad University - Najaf Abad Branch, Isfahan, Iran in 2007. He received his M.Sc. degree in power electric engineering from University of Kashan, Isfahan, Iran in 2012. He's currently pursuing Ph.D. degree of power systems engineering at University of Kashan, Isfahan, Iran. His employment experience included the Iranian oil pipeline and telecommunication company, Tehran, Iran. His special fields of interest included microgrids and distribution systems.

Mohsen Rahimi is an Associate Professor at University of Kashan, Kashan, Iran. He received the B.Sc. degree in electrical engineering in 2001 from Isfahan University of Technology, Isfahan, Iran. He obtained both his M.Sc. and Ph.D. degrees in electrical engineering from Sharif University of Technology (SUT), Tehran, Iran, in 2003 and 2011, respectively.He worked for Saba Niroo, a wind turbine manufacturing company, during 2010–2011. His current research interests include modeling, control and stability analysis of power system dynamics with particular interest in control of grid-connected wind turbines, renewable energy sources, distributed generations, and AC/DC microgrids.

How to cite this paper:
Z. Dehghani Arani, S.A. Taher, M.H. Karimi, M. Rahimi, "Coordinated Model Predictive DC-Link Voltage, Current, and Electromagnetic Torque Control of Wind Turbine with DFIG under Grid Faults," Journal of Electrical and Computer Engineering Innovations, 8(2): 201-218, 2020.

DOI: 10.22061/JECEI.2020.7031.353

URL: http://jecei.sru.ac.ir/article_1465.html

Journal of Electrical and Computer Engineering Innovations
(JECEI)
Journal homepage: http://www.jecei.sru.ac.ir

Research paper

A Novel Hybrid Genetic Algorithm to Predict Students' Academic Performance

Y. Rohani, Z. Torabi*, S. Kianian

Faculty of Computer Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran.

Article Info

Article History:
Received 15 September 2019
Reviewed 16 November 2019
Revised 23 January 2020
Accepted 23 May 2020

Keywords:
Classification
Educational data mining
Simulated annealing algorithm
Genetic algorithm
Educational performance prediction

*Corresponding Author's Email Address:
z.torabi@sru.ac.ir

Abstract

Background: Prediction of students' academic performance is essential for systems emphasizing students' greater success. The results can largely lead to increase in the quality of the educating and learning. Through the application of data mining, useful and innovative patterns can be extracted from the educational data.

Methods: In this paper, a new metaheuristic algorithm, combination of simulated annealing and genetic algorithms, is proposed for predicting students' academic performance in educational data mining. Although metaheuristic algorithms are one of the best options for discovering the hidden relationships between data in data science, they do not separately perform well in accurate prediction of students' academic performance. Therefore, the proposed method integrates the advantages of both genetic and simulated annealing algorithms. The genetic algorithm is applied to explore new solutions, while simulated annealing is used to increase the exploitation power. By using this combination, the proposed algorithm has been able to predict the students' academic performance with high accuracy.

Results: The efficiency of the proposed algorithm is evaluated on five different educational data sets, including two data sets of students of Shahid Rajaee University of Tehran and three online educational data sets. Our experimental results show $1.09\% - 24.39\%$ and $0.29\% - 6.57\%$ accuracy improvement of the proposed algorithm in comparison to the four similar metaheuristic and five popular classification methods respectively.

Introduction

Educational data mining (EDM) refers to the application of data mining methods on data sets, obtained from educational centers or web sites, to extract useful and novel information from data. Early prediction of the students' academic performance at the end of a training course is one of the challenges in educational centers. Today the coronavirus has forced many educational centers to e-learning. It seems they continue online education even after the pandemic ends however some studies [1]-[5] demonstrate higher dropout rates in web-based courses than traditional education ones which increase the importance of EDM. Usage of EDM can prevent financial and psychological consequences which are caused by failure of students, therefore various algorithms and methods have been used for this purpose. Metaheuristic algorithms, which are widely used to solve optimization problems, also have been used by researchers to analyze data and discover hidden patterns in a large data set [6]-[11].

To solve optimization problems effectively, several metaheuristic methods have been proposed such as: particle swarm optimization algorithm [12], artificial bee colony algorithm [13], differential evolution algorithm

[14], firefly algorithm [15] and earthworm optimization algorithm [16]. In these population-based optimization algorithms, the two concepts of "Exploration" and "Exploitation" have received much attention. Exploration refers to the ability to search various unknown areas in the space of a solution in order to find the probable solutions, while exploitation refers to the ability of an algorithm to improve existent solutions, to increase their quality [17].

Genetic algorithm (GA) as a population-based metaheuristic algorithm has both exploration and exploitation strategy and has been widely used in solving optimization problems [8], [18], [19]. Crossover and mutation operators focus more on exploring the problem space, while selection operator focuses on exploiting existing solutions. Thus, in GA algorithm the power of exploration is high, but the power of exploitation is not. In this paper, we proposed a new GA algorithm using simulated annealing (SA) approach to achieve the optimal combination of exploration and exploitation. SA is a single-solution based algorithm that starts with a solution and tries to enhance it. In other words, SA has only exploitation strategy and could strengthen the power of the genetic algorithm. By using this combination, the proposed algorithm has been able to produce high-quality solutions. Therefore, the algorithm can predict students' academic performance more accurately in comparison with other well-known classification methods.

Based on the above explanations, the contributions of this paper could be summarized as follow:

- Proposing an efficient algorithm for classification problems in educational data sets, which is based on combination of GA and SA.
- Proposing a new mutation operator for GA.
- Implementing the proposed algorithm on different educational data sets which leads to the highest prediction accuracy in comparison with the best classification methods.

The rest of this paper is organized as follows. Some recent works on educational data mining and combining metaheuristic algorithms are explained briefly in related work section. In genetic algorithm and simulated annealing algorithm sections, the concept and definitions of the GA and SA algorithms are described respectively. The proposed method is discussed thoroughly in proposed method section. The implementation details and comparison of the proposed work and other common methods are provided in results and discussion section. Finally, conclusion of this paper is presented.

Related Work

In EDM, statistics, machine learning, and data mining techniques are utilized to analyze data collected during teaching and learning [20] in order to find appropriate solutions to educational research. Such solutions are used by both instructors and students to improve teaching and learning. Early prediction of students' learning outcomes is one of the main concerns in EDM which can be categorized in two classes [21]:

- Predicting students' score in a specific course or score point average (GPA).
- Predicting students' academic performance (pass/fail) or dropout.

Unlike some works such as prediction of students' engagement [22] or prediction of slow learners [23], most of the studies about prediction models fall into this category. Several methods have been recently utilized in the prediction of students' grade which is a continuous value [24], [25]. In addition, predicting students' academic performance (pass/fail) or dropout, has been receiving significant attention. In this case, the main goal is to construct a learning model that predicts whether a student will pass/fail or dropout/complete a course.

Students' academic performance or grade depends on several factors such as background characteristics, previous scores, teaching and learning approaches, relationships between student-student, student-teacher, and student-content [25], [26]. Researches in EDM propose prediction models through some of these factors. For example, [27] with consideration of background characteristics and teaching features, presented a study to investigate whether the performance of teachers can predict students' academic performance. Table 1 summarizes recent researches proposed to predict students' outcome. Nevertheless, some works from each category are briefly described here.

For the prediction of final grades, the authors of [28] applied popular algorithms such as Ordinary Least Squares (OLS), Support Vector Machine (SVM), Classification and Regression Tree (CART), k-Nearest Neighbor (kNN), Random Forest (RF) and AdaBoost R2 where SVM reaches the best result. In similar work [29], Kostopoulos proposed a multi-scheme semi-supervised regression approach (MSSRA) using three different k-NN algorithms regressors. The prediction model is based on features such as background characteristics, academic performance and interactions within the learning platform where the results showed the superiority of the MSSRA in comparison with other regression methods.

To achieve accurate prediction of students' dropout, Mubarak et al. [30] extracted significant features from students' weekly interaction with course content. They presented two models based on Logistic Regression (LR) and Input-Output Hidden Markov Model (IOHMM), which have better prediction accuracy in comparison

with baseline of machine learning models. Moreover, with offering instructors' intervention methods, they reduced the rate of dropout. In similar study [31], Burgos et al. conducted a study over the scores of 100 students for several distance learning courses where Logistic Regression models were used for students' dropout prediction. With the usage of this result, they designed a tutoring action plan reducing the dropout rate by 14%. With the aim of discovering patterns that motivate students to drop out, Sarrah et al. [25] developed the Bayesian Profile Regression (BPR) method to identify students who are more likely to drop out. Due to the performance, motivation and resilience of students, this technique draws the profile of students at high risk of academic failure.

In order to predict students' academic performance, Chui et al. [32] proposed a reduced training vector-based support vector machine (RTV-SVM) algorithm. In their research three classes are defined, namely, pass, marginal, and fail.

The authors showed that the RTV-SVM has reached accuracy of 91.2%. Moreover, in large database, the RTV-SVM can be adopted to reduce the training time. In another study [33], the authors addressed high students' failure rates in introductory programming courses. Therefore, several educational data mining techniques were used for prediction of students' academic failure on two data sets including personal and educational information about 262 students from distance education and 161 students from on-campus. They also analyzed the impact of preprocessing and fine-tuning of input parameters to increase the prediction accuracy where the results showed that SVM reaches the best performance. All these studies, manipulate data sets of students' learning behavior, activities, and interactions stored in files and databases. One of the main concerns of such works is acquiring the highest prediction accuracy. Nevertheless, they applied different methods to obtain notable results.

Table 1: Research on prediction of students' learning outcomes

Ref.	Methods	Aim	Type
[28]	OLS, SVM, CART, kNN, RF , AdaBoost R2	Prediction of students' grade	Regression course grades
[43]	GA, Quadratic Bayesian Classifier, kNN, Parzen-window, Multi-Layer Perceptron, Decision Tree.	Prediction of students' grade	Regression course grades
[29]	MSSRA	Prediction of students' grade	Regression course grades
[44]	J48, REPTree	Prediction of students' grade	Regression course grades
[32]	RTV-SVM	Prediction of students' academic performance	Multi-class classification
[25]	BPR	Prediction of students' academic performance	Binary classification
[45]	Gradient Boosting Machine	Prediction of students' academic performance	Binary classification
[33]	Naive Bayes, Decision Tree, SVM, Neural Network	Prediction of students' academic performance	Binary classification
[46]	Decision Tree, Naive Bayes, Neural Network, SVM,kNN	Prediction of students' academic performance	Binary classification
[47]	Regression, Decision Tree	Prediction of students' academic performance	Binary classification
[48]	Ensemble model of Decision Tree, Gradient Boost algorithm and Naive Bayes	Prediction of students' academic performance	Binary classification
[49]	Naive Bayes, J48, RF, Naive Bayes Multiple Nominal, K-star and IBk	Prediction of students' academic performance	Multi-class classification
[31]	LR	Prediction of students' dropout	Binary classification
[30]	LR, IOHMM	Prediction of students' dropout	Binary classification
[50]	JRip, OneR, PART and Ridor	Prediction of students' dropout	Binary classification
[51]	Improved Decision Tree algorithm based on ID3	Prediction of students' dropout	Binary classification
[52]	Multilayer Perception, Naive Bayes, SMO, J48, REPTree	Prediction of slow learners	Binary classification
[53]	Decision Tree, J48, Naive Bayes, CART, JRIP Decision Rules, Gradient Boosting Trees,	Prediction of students' engagement	Binary classification

In this paper, we focus on two popular metaheuristic algorithms, namely SA and GA to obtain an accurate binary classifier in EDM to predict the students' academic performance. In recent years, the combination of metaheuristic algorithms has been used by numerous researchers in the field of optimization [6]. Moreover, hybrid metaheuristic algorithms have shown superior performance in solving many practical or academic problems [7]. In the following, several works on application of hybrid metaheuristic algorithms in different problems are described.

SA is used in numerous hybrid metaheuristic algorithms. In [9], Martin and Otto introduced a hybrid algorithm between the Markov chain and SA, in which the Markov chain is allocated just to detect local optimizations. With combination of two Tabu search algorithms and SA, Lenin et al. [10] proposed a new way to solve the reaction power problem.

In [11], similar combination proposed for symmetric traveling salesman problem. Wang et al. [34] proposed a new hybrid SA for scheduling in dual-resource cellular manufacturing system. In [35], to achieve an automatic diabetic retinopathy screening system, a new hybrid algorithm was proposed by using SA and ensemble bagging classifier. For feature selection, different hybrid metaheuristic algorithms were developed such as combination of local search operations and GA [8], combination of whale algorithm and SA [36]. Several combinations of SA and GA have been proposed in the literature for optimization of signal timing, navigation and routing in the supply chain, thermal structure problem, and flow shop scheduling [37]-[42].

Genetic Algorithm

Genetic algorithm is a metaheuristic algorithm inspired by the principle of the natural selection and natural genetics [18]. GA represents the solutions in the form of chromosomes and the fitness of the chromosomes is evaluated by fitness function which is created according to the objective function of the optimization problem. A collection of chromosomes is called a population where initially a random population is created. Individual solutions are selected to be parents for generating a new population through their fitness values, where fitter solutions are typically more likely to be selected. A pair of parent solutions creates new children by crossover and mutation operators where new solutions typically share many features of their parents. Mutation operator is used to avoid getting stuck in the optimal local trap. The process of selection, crossover, and mutation operators improves the population and continues until a new population of appropriate size is generated. The best solution in the last population is returned as the best approximation of the global optimum. Compared to the traditional metaheuristic methods, GA is well converged due to the adaptation of the biological evolution model [54]. Algorithm 1 shows the pseudo code of GA.

Algorithm 1: Genetic Algorithm

```
begin
  Generate initial population
  repeat
    Evaluate the individual solutions
    Select pairs of best-ranking individuals
    Apply crossover operator
    Apply mutation operator
  until terminating condition is not met
end
```

Simulated Annealing Algorithm

SA proposed by Kirkpatrick et al. [55] is a metaheuristic method inspired by the annealing procedure used in metallurgy, suitable in solving complex optimization problems. SA starts with a randomly generated solution in high temperature. In each iteration, a neighbor solution is generated according to a predefined neighborhood structure and evaluated using a fitness function. If the new solution is better than the original one, it is accepted, otherwise it can still be accepted with probability $e^{-\frac{\theta}{k_B T}}$, where θ is the difference between the fitness of the current solution and the generated neighbor, k_B is Boltzmann's constant and T is the current temperature [36]. By accepting worse solutions, SA can avoid being trapped on local optimum. The parameter T is gradually decreased by a cooling function as SA proceeds until the termination condition is met [56]. Algorithm 2 shows the pseudo code of SA algorithm.

Algorithm 2: Simulated Annealing Algorithm

```
begin
  Initialize solution x, temperature T
  repeat
    repeat
      Create neighborhood solution y
      if θ ≤ 0 then accept y
      otherwise accept y by the probability e^(-θ/(k_B T))
    until inner-loop terminating condition is not met
    Decreasing T gradually
  until outer-loop terminating condition is not met
end
```

Proposed Method

In this paper, the G-SA algorithm, which is a new hybrid of GA and SA algorithms, is presented to predict

students' academic performance. The framework of the G-SA algorithm is shown in Fig. 1. As the figure shows, an initial population of states is generated first. After generating the initial population, one randomly chosen state, namely current state, is utilized to crossover with the best state of the population (i.e. Gbest) where offspring1 and offspring2 are produced. On the other hand, an offspring3 is created by the mutation operation. The best of these three offsprings is compared to the current state. If the best offspring is better, it would replace the current state. Otherwise it would replace with probability $e^{-\frac{\theta}{k_B T}}$. This process continues until the new population would be completed with n new states where n is the size of population. After that, based on the current temperature it is decided to generate a new population or stop generating new population. The best state of final population is added to the rule set and all the states from the training set that can be evaluated by this rule are removed from the training set.

The algorithm runs again with the remaining states of the training set to obtain another rule. This continues until the size of the training set reaches a certain threshold (10% of the training set). In the end, a set of rules is obtained that is used for prediction. Algorithm 3 shows how the G-SA works.

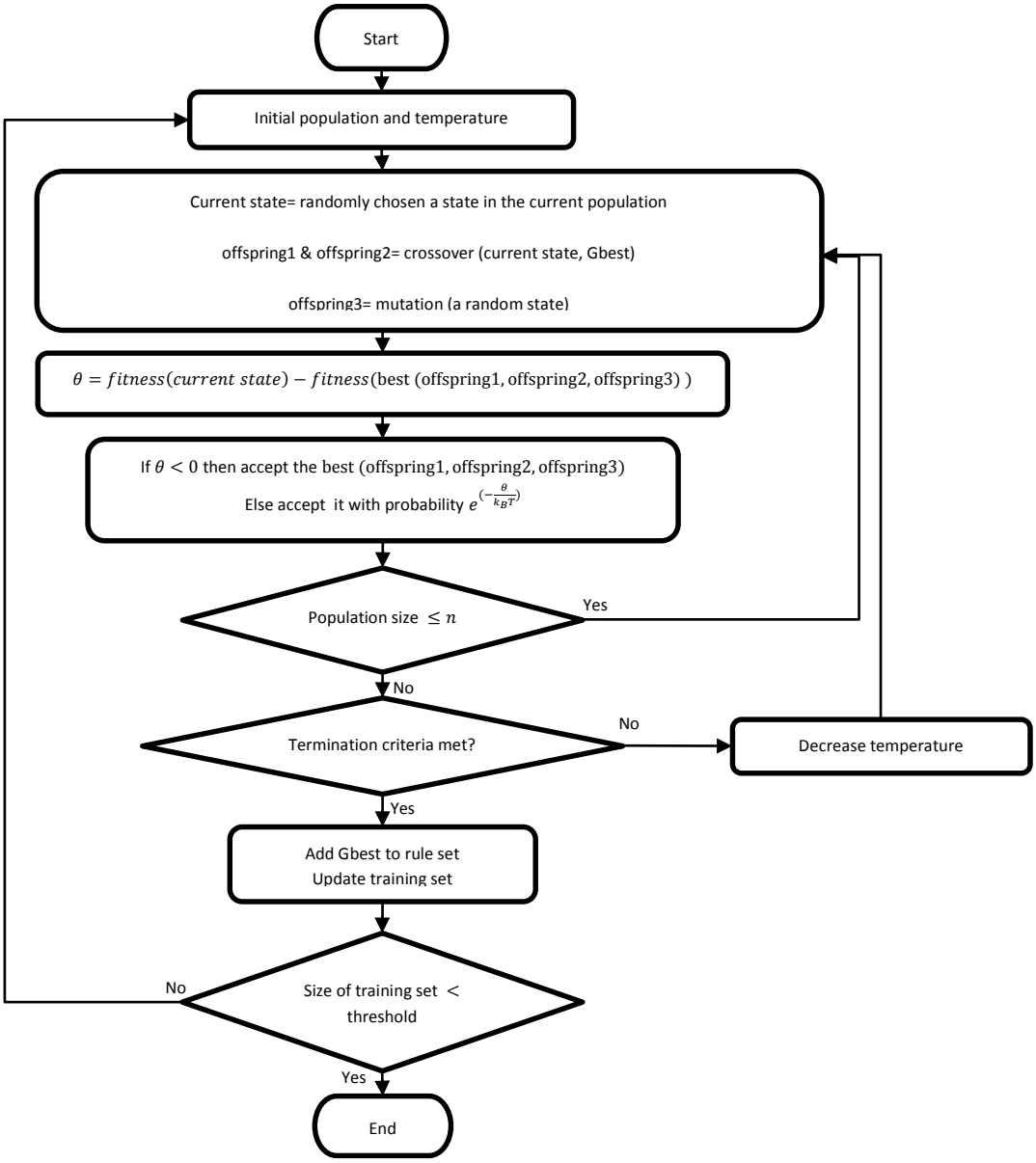


Fig. 1: Framework of the G-SA algorithm.

Algorithm 3: G-SA Algorithm

```
begin
 repeat
   Initialize first population and T
   Find the best state obtained so far, which has highest fitness value (i.e.Gbest)
   repeat
     repeat
       (Crossover)
         Select random state u as current state
         for each feature index j in u
           Select randomly bit RN between [0, 1]
           if RN > 0.6 then
             Offspring1_j = Gbest_j
             Offspring2_j = u_j
           else
             Offspring1_j = u_j
             Offspring2_j = Gbest_j
       (Mutation)
         Select random state v
         Offspring3 = v
         for each feature index j in v
           Select randomly bit RN between [0, 1]
           if RN <= 0.1 then
             Update Offspring3_j value using Equation 1
       new state = best (offspring1, offspring2, offspring3)
       θ = fitness(u) − fitness(new state)
       if θ ≤ 0 then Accept new state
       otherwise Accept it by the probability e^(−θ/(k_B T))
       (Update Gbest)
         Find the best state obtained so far, which has highest fitness value
     until population size reaches n
     Decrease T gradually
   until outer-loop terminating condition is not met
   Add Gbest as a rule to rule set
   Update training set
 until the size of the training set reaches threshold
end
```

A. Problem formulation and initial population

The aim of the G-SA algorithm is to predict the students' academic performance (pass/fail) using some background characteristics and previous scores.

Let $X = [x_1, x_2, \dots, x_n]$ be pattern of the initial population where x_i is a chromosome that represents a candidate solution which is modeled as an array of genes. In educational data sets, genes of a chromosome are dependent on a given data set and may contain background characteristics, previous scores, teaching and learning approaches, relationships between student-student, student-teacher, and student-content. Last gene in each chromosome indicates academic performance. Operators of GA in the proposed method are applied to these chromosomes. Here, since the proposed algorithm is a combination of SA and GA, these chromosomes are called states.

For more clarification, Fig. 2 shows an example of a state in our problem modeling for Math dataset explained in Section 6-A. In this figure, the first 10 features are background characteristics of student such as age, sex, parents' jobs and educations. Next 3 features are scores of three exams and last feature represents the success or failure of student and must be predicted.

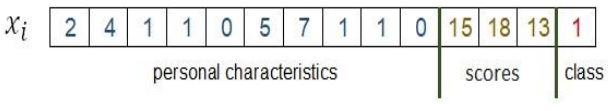


Fig. 2: An example of a state.

B. Crossover and Mutation

One of the important points that can directly affect the accuracy and convergence speed of the algorithm is the neighborhood selection strategy. In basic SA, some operators such as insertion, reversion and swap are used

to produce a neighborhood state, which sometimes does not result in a global optimal solution. Crossover and mutation operators in GA can be good options for generating new neighborhood states. The quality of the new states generated by crossover and mutation operators depends on the fitness of the parents. In this study, the best state in the population named Gbest is used in crossover operator.

Gbest and a randomly chosen state (i.e. current state) are combined under crossover operation to generate offspring1 and offspring2. The G-SA uses the crossover operation [57] implemented by generating a vector of random numbers in the range [0, 1] having the same length as the parents. In each feature of a current state, if the value of the random vector is below 0.6, offspring1 takes this feature from Gbest, otherwise from a current state. If the value of the random vector is above 0.6, offspring2 takes this feature from a current state, otherwise from Gbest.

One-point or two-point uniform crossover operators are usually used in GA where parents are randomly selected based on their fitness from a large search space. The crossover operator of [57] by deciding for each feature independently led to accurate classifier [57]. In addition, one of the parents in the crossover operator of [57] is Gbest, increasing the fitness of the produced offsprings and convergence speed.

After performing the crossover and producing two offsprings, offspring3 is generated by the mutation operator. In mutation operator, another random state is selected and a random vector is generated in the range [0, 1] having the same length as the random state. If the value of each feature of random vector is less than or equal to 0.1 then the corresponding feature of random state would change according to (1), where i is a number of feature, $\emptyset_i$ is a random number from a normal distribution with mean 0 and standard deviation $sigma$. For each feature, $sigma$ is obtained according to (2), where $Varmax_i$ and $Varmin_i$ are the upper and lower bounds of the *i*-th feature respectively.

$$v_i = x_i + \emptyset_i * sigma_i \quad (1)$$

$$sigma_i = 0.1 * (Varmax_i - Varmin_i) \quad (2)$$

Fig. 3, shows an example of how a mutation operator works. As can be seen in this example, only the fourth feature of the random vector is less than 0.1. Therefore, by changing it in random state according to (1), a new offspring is produced that differs from the random state only in one feature.

In the conventional mutation operator, a new offspring is generated by changing only one feature of parent state which leads to the offspring different from the parent in only one feature. To avoid getting stuck in local optimum, the proposed mutation operator in this study, can change more than one features increasing the mutation power. Therefore, the proposed mutation operator by creating offspring with more distance from the parent state, prevents the algorithm trapping into local optimum.

Fig. 3: Mutation operator in G-SA algorithm.

C. *Fitness function*

In this paper, (3), is used to evaluate the states in population. Each state in the training set is considered as a rule for prediction. All the states in the training set are compared to this rule, and the values of the two TP and TN criteria are specified.

$$f(x_i) = \frac{TP + TN}{size(Training\ Set)} \quad (3)$$

The TP criterion contains the number of samples that are correctly predicted as positive and TN contains the number of samples that are correctly predicted as negative.

Results and Discussion

To indicate the superiority of the G-SA algorithm, this section presents and analyzes the results of the proposed algorithm and other state-of-the-art classification methods. The G-SA algorithm has been compared to five well-known classification algorithms in data mining, including Decision Tree J48, Naive Bayes (NB), Multilayer Perceptron (MP), LR, and SVM. Furthermore, it has been compared to the four following metaheuristic methods.

- Basic ABC algorithm [58].
- GBC algorithm [57].
- Basic SA algorithm [55].
- SA-GA algorithm [41].

In the following, datasets which are used for evaluation are described and then parameters of algorithms are estimated. Finally, the G-SA algorithm results are compared with other classification methods.

A. Data Sets

In order to evaluate the proposed algorithm for prediction of student performance, five different educational data sets including S5f, E-circ, Math, Port, and Deeds have been used.

The detailed descriptions of these five data sets are presented in Table 2 with respect to the number of classes, samples, and features.

S5f and E-circ data sets have been obtained from the database of computer engineering students at Shahid Rajaee University of Tehran during the years of 2011-14. S5f data set includes the students' average grades in the first five semesters.

The purpose is to predict the performance of students in the fifth semester using the average grades of the first four semesters. E-circ data set is compiled to predict students' performance in Electric Circuits course which is one of the most challenging courses using the scores of Discrete Structures, Physics, Mathematics, Differential Equations and Logic Circuits.

The other three data sets were compiled by researchers in the field of educational data mining to predict the educational academic status of students and can be accessed online on the UCI[1] site. Math and Port data sets related to the students of Portuguese school, whose data were gained in two courses of Mathematics and Portuguese Language respectively and have 33 identical features [59].

Since the large number of features increases the execution time of the algorithm greatly, first feature selection operation was applied to these two data sets. For this purpose, with the Information Gain method in WEKA, data mining tool, these data sets were analyzed and among the 33 features, top 10 features were selected for each of the Math and Port data sets.

Deeds data set is related to Logic Circuits course at the University of Genoa, Italy, and includes scores of 16 tests out of 100 samples, with the final test score being considered as a class [60]. In Deeds data set, among its features, six characteristics were selected to predict the class.

Table 2: Statistics of educational data sets

Data sets	Number of features	number of samples	Class Fail	Class Pass
S5f	5	128	10	118
E-circ	8	255	105	150
Math	11	395	130	265
Port	10	649	100	549
Deeds	6	114	63	51

[1] https://archive.ics.uci.edu/ml/index.php

The features of S5f, E-circ and Deeds data sets are some exam scores in different courses. In the Math and Port data sets, in addition to the three test scores, they also include some personal characteristics of students. Table 3 shows the features of the Math and Port data sets. Some of these features are used in the math and some in the port data set.

Table 3: Features of the math and Port database

Attribute	Description (Domain)
age	student's age (numeric: from 15 to 22)
school	student's school (binary: Gabriel Pereira or Mousinho da Silveira)
Pstatus	parent's cohabitation status (binary: living together or apart)
Medu	mother's education (numeric: from 0 to 4)
famsize	family size (binary: < 3 or > 3)
famrel	quality of family relationships (numeric: from 1 : very bad to 5 : excellent)
reason	reason to choose this school (nominal: close to home, school reputation, course preference or other)
studytime	weekly study time (numeric: 1 : < 2 hours, 2 : 2 to 5 hours, 3 : 5 to 10 hours or 4 : > 10 hours)
failures	number of past class failures (numeric: n if 1 : n < 3, else 4)
schoolsup	extra educational school support (binary: yes or no)
famsup	family educational support (binary: yes or no)
higher	wants to take higher education (binary: yes or no)
freetime	free time after school (numeric: from 1 : very low to 5 : very high)
goout	going out with friends (numeric: from 1 : very low to 5 : very high)
G1	first period grade (numeric: from 0 to 20)
G2	second period grade (numeric: from 0 to 20)
G3	final grade (Class – Fail or Pass)

B. Parameter settings for the algorithms

In this paper, MATLAB has been used for implementation of the algorithms with 10-fold cross validation method for all of the algorithms. To attain a fair comparison, we ran each algorithm with different parameters several times to find appropriate initialization values for the best results.

In ABC-based algorithms (i.e. ABC and GBC) the best results in five data sets are obtained by setting the maximum number of iterations to 400 and the colony size to 40 and both the number of food source and the rate of food source abandonment to 20.

In the implementation of SA-based algorithms (i.e. SA and SA-GA) the iteration of the main and sub loop set to 1000 and 10 respectively. Also, the initial population, temperature and the rate of temperature reduction, set to 20, 10 and 0.95 respectively. Due to the random nature of the initial population selection, each of the algorithms runs 50 times, and the average of accuracy was considered as the accuracy of the algorithm.

Population size of the G-SA algorithm set to 1000. Accuracy of the proposed method with different values for temperature reduction $(Alpha)$ and initial temperature (T_0) on the educational data sets are shown in Table 4. As can be seen, in the four data sets, the G-SA algorithm with $T_0 = 10$ and $Alpha = 0.95$ has achieved the highest accuracy, so it is considered as the initial settings for G-SA algorithm.

Table 4: The classification accuracy performance of the G-SA algorithm with different T_0 and $Alpha$

parameters	S5f	E-circ	Port	Math	Deeds
T_0=10 ; *Alpha*=0.99	95.62	79.43	94.13	**93.60**	93.62
Alpha=0.95	**96.30**	**80.29**	**94.84**	93.46	**94.49**
Alpha=0.9	95.18	79.22	94.35	93.32	93.21
Alpha=0.8	95.11	78.17	93.75	92.77	92.31
Alpha=0.75	94.43	78.38	93.63	92.41	93.15
T_0=20 ; *Alpha*=0.99	94.26	78.35	93.56	93.29	92.99
Alpha=0.95	95.33	79.91	93.80	93.14	93.82
Alpha=0.9	93.85	77.69	92.62	93.10	91.53
Alpha=0.8	94.27	77.33	93.32	91.33	90.23
Alpha=0.75	92.51	76.79	92.17	92.11	89.22

C. Experimental results of comparisons with other classification methods

The evaluation criterion for comparing the G-SA and other algorithms are accuracy and F-measure. Accuracy is computed by dividing the set of correct predictions by the sum of all predictions according to the (4).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

The variables used in the Equation 4 are as follow:

TP: The number of true positive samples that are correctly predicted as positive.

TN: The number of true negative samples that are correctly predicted as negative.

FP: The number of true negative samples that are incorrectly predicted as positive.

FN: The number of true positive samples that are incorrectly predicted as negative.

In F-measure criterion, the incorrect prediction rate along with the correct prediction rate, are used to evaluate the efficiency of the algorithms. Equation (5) shows the F-measure criterion, in which the two criteria Precision and Recall are obtained according to (6) and (7), respectively.

$$F - measure = \frac{2 * Precision * Recall}{Precision + Recall} \quad (5)$$

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

Table 5 shows the accuracy of the G-SA algorithm along with other relevant metaheuristic algorithms.

The numbers marked in blue are the highest accuracy in each data set related to the G-SA method.

The G-SA algorithm has improved accuracy compared to the basic ABC algorithm in the data set of Math and Deeds by 19.01% and 24.39%, respectively.

The minimum improvement in accuracy by the proposed method compared to the basic ABC algorithm has occurred in the data set of the Port which is 4.69%.

Maximum and minimum accuracy improvements compared to the GBC are 19.69% in the data set of E-circ and 5.15% in the data set of Port.

In addition, the maximum accuracy improvement of the G-SA algorithm compared to the SA and SA-GA methods are related to the data set of Deeds, which are 6.61% and 3.15%, respectively and the minimum accuracy improvement over the SA and SA-GA methods are 1.62% and 1.09% in the data set of Port.

Table 5: Accuracy performance of the proposed method, ABC, GBC, SA, and SA-GA

Data Sets	ABC	GBC	SA	SA-GA	G-SA
S5f	91.33	86.05	93.81	94.22	**96.30**
E-circ	61.28	60.60	77.13	78.48	**80.29**
Port	90.15	89.69	93.22	93.75	**94.84**
Math	76.75	77.34	89.00	92.10	**93.52**
Deeds	70.10	74.39	87.88	91.34	**94.49**

The accuracy of the G-SA algorithm, along with other popular classification methods, can be seen in Table 6. The proposed algorithm has the highest accuracy which marked in blue in each data set. The G-SA algorithm in different data sets, compared to the J48, has improved the accuracy from 1.12% to 6.57%. The highest and lowest accuracy improvements of the G-SA algorithm in comparison with NB are 5.68% and 0.64% in the Math and Deeds data sets respectively. The maximum accuracy improvement compared to SVM and MP are 4.66% and 5.68% in Math data set. The minimum accuracy improvement of the G-SA in Table 6 is 0.29% and related to LR in the data set of E-circ.

Table 6: Accuracy performance of the G-SA and five well-known classification methods

Data Sets	J48	NB	LR	SVM	MP	G-SA
S5f	94.53	93.75	94.53	92.97	93.75	**96.30**
E-circ	73.72	75.68	80.00	79.60	75.29	**80.29**
Port	93.62	91.83	92.91	91.52	90.91	**94.84**
Math	92.40	87.84	91.49	88.86	87.84	**93.52**
Deeds	88.59	93.85	92.98	93.85	93.85	**94.49**

The results of the proposed method with F-measure criterion in comparison with metaheuristic algorithms and conventional classification methods are shown in Tables 7 and 8, respectively. As can be seen, the results of the F-measure criterion also show better performance in comparison with other methods.

For better representation, Fig. 4, and Fig. 5, show the accuracy and F-measure results of the G-SA algorithm respectively, along with relevant metaheuristic algorithms and conventional classification methods. The results show that combination of two metaheuristic algorithms in G-SA leads to more accurate results.

Table 7: F-measure criterion of the proposed method, ABC, GBC, SA, and SA-GA

Data Sets	ABC	GBC	SA	SA-GA	G-SA
S5f	0.92	0.88	0.93	0.95	**0.96**
E-circ	0.63	0.61	0.77	0.78	**0.81**
Port	0.90	0.90	0.93	0.94	**0.97**
Math	0.78	0.79	0.90	0.92	**0.94**
Deeds	0.71	0.74	0.88	0.93	**0.95**

It can be noted that in most cases, the accuracy of each of the metaheuristic algorithms (i.e. ABC, GBC, SA, and SA-GA) is lower than the conventional classification methods, however, in G-SA algorithm after combining SA and GA properly, accuracy improved significantly.

Table 8: F-measure criterion of the proposed method and five well-known methods in classification

Data Sets	J48	NB	LR	SVM	MP	G-SA
S5f	0.94	0.93	0.94	0.90	0.93	**0.96**
E-circ	0.73	0.76	0.80	0.79	0.75	**0.81**
Port	0.93	0.92	0.92	0.90	0.91	**0.97**
Math	0.92	0.88	0.91	0.88	0.87	**0.94**
Deeds	0.89	0.94	0.93	0.94	0.94	**0.95**

Experimental results confirm the good balance between the power of exploration and exploitation of the proposed algorithm.

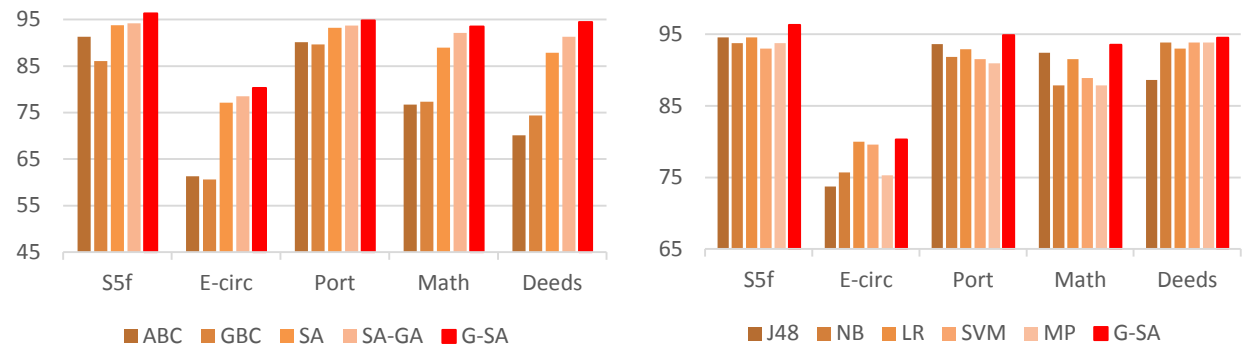
The crossover operator in the proposed algorithm utilized from the best global state generates high fitness states.

Such crossover operator sometimes may cause to trap the algorithm in local optimum.

The proposed method by introducing the new mutation operator prevents trapping in local optimum.

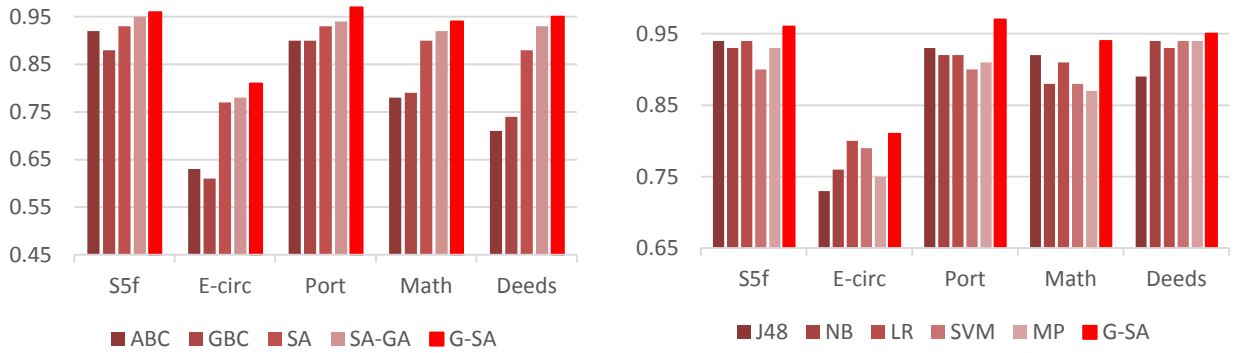
In the new mutation operator, instead of changing only one feature, more than one features may be changed, which increases the mutation power.

Due to the methaheuristic strategy and random nature of the initial population selection from the training set, increasing the volume of data not only would not limit the algorithm, but also can probably increase the accuracy as can be seen in the experimental results.



(a) (b)

Fig. 4: Diagram of the accuracy performance of the proposed method and nine other classifiers.



(a) (b)

Fig. 5: Diagram of the F-measure criterion of the proposed method and nine other classifiers.

Conclusion

In this paper, a new method called G-SA, is presented for prediction of students' academic performance during educational courses, which can be used to prevent possible failures of students. The G-SA algorithm uses the advantages of both simulated annealing and genetic algorithms by combining them. By relying on the best global solution in the proposed algorithm, crossover and mutation operators produce stronger neighbors and ultimately lead to better solutions. The combination of the two algorithms also balances the power of exploration and exploitation in the proposed algorithm, which has not only helped to speed convergence, but has also been able to get rid of local optimum. Experimental results from the implementation of the G-SA algorithm show that the proposed algorithm improves accuracy performance from 1.09% to 24.39% compared to other metaheuristic comparison methods and from 0.29% to 6.57% compared to well-known conventional classification methods.

Author Contributions

This paper is the result of Y. Rohani's MSc thesis supervised by Z. Torabi, and S. Kianian.

Conflict of Interest

The authors declare that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or

falsification, double publication and/or submission, and redundancy have been completely observed by the authors.

Abbreviations

GA	Genetic Algorithm
SA	Simulated Annealing
G-SA	Genetic Simulated Annealing Algorithm
Gbest	Global best states
J48	Decision Tree J48
NB	Naive Bayes
LR	Logistic Regression
SVM	Support Vector Machine
MP	Multilayer Perceptron
n	Size of Population
k_B	Boltzmann's Constant
T	Current Temperature
S5f	Student performance data set in the fifth semester
E-circ	Student performance data set in Electric Circuits course
Math	Student performance data set in Mathematics course
Port	Student performance data set in Portuguese language course
Deeds	Student performance data set in Digital Electronics Education and Design Suite

References

[1] E. Black, K. Dawson, J. Priem, "Data for free: Using LMS activity logs to measure community in online courses," The Internet and Higher Education, 11(2): 65-70, 2008.

[2] C. Chen, Y. Chen, C. Liu, "Learning performance assessment approach using web-based learning portfolios for e-learning systems," IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 37(6): 1349-1359, 2007.

[3] J. Hung, K. Zhang, "Revealing online learning behaviors and activity patterns and making predictions with data mining techniques in online teaching," MERLOT Journal of Online Learning and Teaching, 4(4): 426-437, 2008.

[4] M. Roblyer, L. Davis, S. Mills, J. Marshall, L. Pape, "Toward practical procedures for predicting and promoting success in virtual school students," The Amer. Jrnl. of Distance Education, 22(2): 90-109, 2008.

[5] A. Juan, T. Daradoumis, J. Faulin, F. Xhafa, "Developing an information system for monitoring student's activity in online collaborative learning," in Proc. International Conference on Complex, Intelligent and Software Intensive Systems: 270-275, 2008.

[6] E. Talbi, Metaheuristics: from design to implementation. John Wiley & Sons; 2009.

[7] E. Talbi, "A taxonomy of hybrid metaheuristics," Journal of heuristics, 8(5): 541-564, 2002.

[8] I. Oh, J. Lee, B. Moon, "Hybrid genetic algorithms for feature selection," IEEE Transactions on pattern analysis and machine intelligence, 26(11): 1424-37, 2004.

[9] O. Martin, S. Otto, "Combining simulated annealing with local search heuristics. Annals of Operations Research," 63(1): 57-75, 1996.

[10] K. Lenin, B. Reddy, M. Suryakalavathi, "Hybrid Tabu search-simulated annealing method to solve optimal reactive power problem," International Journal of Electrical Power & Energy Systems, 82: 87-91, 2016.

[11] Y. Lin, Z. Bian, X. Liu, "Developing a dynamic neighborhood structure for an adaptive hybrid simulated annealing–tabu search algorithm to solve the symmetrical traveling salesman problem," Applied Soft Computing, 49: 937-52, 2016.

[12] J. Kennedy, R. Eberhart, "Particle swarm optimization," in Proc. of ICNN'95-International Conference on Neural Networks, 4: 1942-1948, 1995.

[13] D. Karaboga, "An idea based on honey bee swarm for numerical optimization," Technical report-tr06, Erciyes university, engineering faculty, computer engineering department; 2005.

[14] Z. Meng, J. Pan, "HARD-DE: Hierarchical archive-based mutation strategy with depth information of evolution for the enhancement of differential evolution on numerical optimization, IEEE Access, 7: 12832-54, 2019.

[15] H. Wang, W. Wang, H. Sun, S. Rahnamayan, "Firefly algorithm with random attraction," International Journal of Bio-Inspired Computation, 8(1): 33-41, 2016.

[16] G. Wang, S. Deb, L. Coelho, "Earthworm optimisation algorithm: a bio-inspired metaheuristic algorithm for global optimisation problems," International Journal of Bio-Inspired Computation, 12(1): 1-22, 2018.

[17] A. Singh, K. Deep, "Exploration–exploitation balance in Artificial Bee Colony algorithm: a critical analysis," Soft Computing, 23(19): 9525-9536, 2019.

[18] H. Braun, "On solving travelling salesman problems by genetic algorithms," in Proc. International Conference on Parallel Problem Solving from Nature: 129-133, 1990.

[19] Y. Deng, Y. Liu, D. Zhou, "An improved genetic algorithm with initial population strategy for symmetric TSP," Mathematical Problems in Engineering, 2015: 1-7, 2015.

[20] C. Romero, S. Ventura, "Data mining in education," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 3(1): 12-27, 2013.

[21] M. Tsiakmaki, G. Kostopoulos, S. Kotsiantis, O. Ragos, "Implementing AutoML in educational data mining for prediction tasks," Applied Sciences, 10(1): 90, 2020.

[22] T. Soffer, A. Cohen, "Students' engagement characteristics predict success and completion of online courses," Journal of Computer Assisted Learning, 35(3): 378-89, 2019.

[23] A. Rajeswari, C. Deisy, "Fuzzy logic based associative classifier for slow learners prediction," Journal of Intelligent & Fuzzy Systems, 36(3): 2691-704, 2019.

[24] R. Morsomme, E. Smirnov, "Conformal Prediction for Students' Grades in a Course Recommender System," In Conformal and Probabilistic Prediction and Applications, 2019: 196-213, 2019.

[25] A. Sarra, L. Fontanella , S. Di Zio, "Identifying students at risk of academic failure within the educational data mining framework," Social Indicators Research, 146(1-2): 41-60, 2019.

[26] A. Agudo-Peregrina, S. Iglesias-Pradas, M. Conde-González, A. Hernández-García, "Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and

online learning," Computers in human behavior, 31: 542-50, 2014.

[27] C. Herodotou, B. Rienties, A. Boroowa, Z. Zdrahal, M. Hlosta, "A large-scale implementation of predictive learning analytics in higher education: the teachers' role and perspective," Educational Technology Research and Development, 67(5): 1273-306, 2019.

[28] P. Strecht, L. Cruz, C. Soares, J. Mendes-Moreira, "A Comparative Study of Classification and Regression Algorithms for Modelling Students' Academic Performance," International Educational Data Mining Society, 2015: 392- 395, 2015.

[29] G. Kostopoulos, S. Kotsiantis, N. Fazakis, G. Koutsonikos, C. Pierrakeas, "A Semi-Supervised Regression Algorithm for Grade Prediction of Students in Distance Learning Courses," International Journal on Artificial Intelligence Tools, 28(04): 1940001, 2019.

[30] A. Mubarak, H. Cao, W. Zhang, "Prediction of students' early dropout based on their interaction logs in online learning environment," Interactive Learning Environments, 20: 1-20, 2020.

[31] C. Burgos, M. Campanario, D. de la Peña, J. Lara, D. Lizcano, M. Martínez, "Data mining for modeling students' performance: A tutoring action plan to prevent academic dropout," Computers & Electrical Engineering, 66: 541-56, 2018.

[32] K. Chui, D. Fung, M. Lytras, M. Lam, "Predicting at-risk university students in a virtual learning environment via a machine learning algorithm," Computers in Human Behavior, 107: 105584, 2020.

[33] E. Costa, B. Fonseca, M. Santana, E. de Araújo, J. Rego, "Evaluating the effectiveness of educational data mining techniques for early prediction of students' academic failure in introductory programming courses," Computers in Human Behavior, 73: 247-56, 2017.

[34] J. Wang, C. Liu, K. Li, "A hybrid simulated annealing for scheduling in dual-resource cellular manufacturing system considering worker movement," Automatika, 60(2): 172-80, 2019.

[35] S. Sreng, N. Maneerat, K. Hamamoto, R. Panjaphongse, "Automated diabetic retinopathy screening system using hybrid simulated annealing and ensemble bagging classifier," Applied Sciences, 8(7): 1198, 2018.

[36] M. Mafarja, S. Mirjalili, "Hybrid whale optimization algorithm with simulated annealing for feature selection," Neurocomputing, 260: 302-12, 2017.

[37] P. Vasant P, "Hybrid simulated annealing and genetic algorithms for industrial production management problems," International Journal of Computational Methods, 7(02): 279-97, 2010.

[38] Z. Li, P. Schonfeld, "Hybrid simulated annealing and genetic algorithm for optimizing arterial signal timings under oversaturated traffic conditions," Journal of advanced transportation, 49(1): 153-70, 2015.

[39] Y. Li, H. Guo, L. Wang, J. Fu, "A hybrid genetic-simulated annealing algorithm for the location-inventory-routing problem considering returns under E-supply chain environment," The Scientific World Journal, 2013: 1-11, 2013.

[40] L. Junghans, N. Darde, "Hybrid single objective genetic algorithm coupled with the simulated annealing optimization method for building optimization," Energy and Buildings, 86: 651-62, 2015.

[41] H. Wei, S. Li, H. Jiang, J. Hu, J. Hu, "Hybrid genetic simulated annealing algorithm for improved flow shop scheduling with make span criterion," Applied Sciences, 8(12): 2621, 2018.

[42] F. Erchiqui, "Application of genetic and simulated annealing algorithms for optimization of infrared heating stage in thermoforming process," Applied Thermal Engineering, 128: 1263-72, 2018.

[43] B. Minaei-Bidgoli, W. Punch, "Using genetic algorithms for data mining optimization in an educational web-based system," in Proc. Genetic and evolutionary computation conference 2003: 2252-2263, 2003.

[44] S. Natek, M. Zwilling, "Student data mining solution–knowledge management system related to higher education institutions," Expert systems with applications, 41(14): 6400-6407, 2014.

[45] I. papadogiannis, V. Poulopoulos, M. Wallace, "A Critical Review of Data Mining for Education: What has been done, what has been learnt and what remains to be seen," International Journal of Educational Research Review, 5(4):353-372, 2020.

[46] A. RİMİ, A. IBRAHİM, O. BAYAT, "Developing Classifier for the Prediction of Students' Performance Using Data Mining Classification Techniques," AURUM Mühendislik Sistemleri ve Mimarlık Dergisi, 4(1):73-91, 2020.

[47] H. Hassan, R. Mohamad, R. Ali, Y. Talib, H. Hsbollah, "Factors Affecting Students' Academic Performance in Higher Education: Evidence from Accountancy Degree Program," International Business Education Journal, 13(1): 1-6, 2020.

[48] P. Kamal, S. Ahuja, "An ensemble-based model for prediction of academic performance of students in undergrad professional course," Journal of Engineering, Design and Technology, 2019.

[49] B. Kapur, N. Ahluwalia, R. Sathyaraj, "Comparative study on marks prediction using data mining and classification algorithms," International Journal of Advanced Research in Computer Science, 8(3): 1-5, 2017.

[50] C. Siebra, R. Santos, N. Lino, "A Self-Adjusting Approach for Temporal Dropout Prediction of E-Learning Students," International Journal of Distance Education Technologies (IJDET), 18(2):19-33, 2020.

[51] S. Sivakumar, S. Venkataraman, R. Selvaraj, "Predictive modeling of student dropout indicators in educational data mining using improved decision tree," Indian Journal of Science and Technology, 9(4): 1-5, 2016.

[52] P. Kaur, M. Singh, G. Josan, "Classification and prediction-based data mining algorithms to predict slow learners in education sector," Procedia Computer Science, 57: 500-8, 2015.

[53] M. Hussain, W. Zhu, W. Zhang, S. Abidi, "Student engagement predictions in an e-learning system and their impact on student course assessment scores," Computational intelligence and neuroscience, 2018..

[54] J. McCall, "Genetic algorithms for modelling and optimisation. Journal of computational and Applied Mathematics," 184(1): 205-22, 2005.

[55] S. Kirkpatrick, C. Gelatt, M. Vecchi, "Optimization by simulated annealing," science, 220(4598): 671-680, 1983.

[56] T. Wu, C. Chang, S. Chung, "A simulated annealing algorithm for manufacturing cell formation problems," Expert Systems with Applications, 34(3): 1609-1617, 2008.

[57] H. Alshamlan, G. Badr, Y. Alohali, "Genetic Bee Colony (GBC) algorithm: A new gene selection method for microarray cancer

classification," Computational biology and chemistry, 56: 49-60, 2015.

[58] D. Karaboga, B. Basturk, "On the performance of artificial bee colony (ABC) algorithm," Applied soft computing, 8(1): 687-97, 2008.

[59] P. Cortez, A. Silva. Using data mining to predict secondary school student performance, in Proc. 5th Annual Future Business Technology Conference: 5-12, 2008.

[60] M. Vahdat, L. Oneto, D. Anguita, M. Funk, M. Rauterberg, "learning analytics approach to correlate the academic achievements of students with interaction data from an educational simulator," in Proc. Design for teaching and learning in a networked world 2015: 352-366, 2015.

Biographies

Yaser Rohani received his BSc in 2002 from the Islamic Azad University of Mashhad, Iran and his MSc in 2020 from Shahid Rajaee Teacher Training University, Tehran, Iran, all in computer engineering. He is currently working as a computer teacher at the Ministry of Education, in Technical Conservatory. His research interests are data mining, optimization and classification algorithms.

Zeinab Torabi received her Ph.D. degree in Computer System Architecture from Shahid Beheshti University, Tehran, Iran, in 2016. She is currently an Assistant Professor of Computer Engineering in Department of Computer Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran. Her research interests include computer arithmetic, residue number system, design and analysis of algorithms.

Sahar kianian received her B.Sc. degree in computer Engineering (2007) from razi University, also M.Sc. and Ph.D. degrees in computer Engineering from Isfahan University (2010 and 2016, respectively). She is an assistant professor of computer engineering at Shahid Rajaee University. Her research interests are the application of algorithms, machine learning and data science to complex networks, focuses on protein interactions, connections of neurons and relationships among people. Applications include disease prediction, drug discovery, event detection and tracking, recommendation system, web mining and social influence mining.

How to cite this paper:
Y. Rohani, Z. Torabi, S. Kianian, "A Novel Hybrid Genetic Algorithm to Predict Students' Academic Performance," Journal of Electrical and Computer Engineering Innovations, 8(2): 219-232, 2020.

DOI: 10.22061/JECEI.2020.7230.373

URL: http://jecei.sru.ac.ir/article_1459.html

Journal of Electrical and Computer Engineering Innovations
(JECEI)
Journal homepage: http://www.jecei.sru.ac.ir

Research paper

An Efficient Configuration for Energy Hub to Peak Reduction Considering Demand Response Using Metaheuristic Automatic Data Clustering

H.Hosseinnejad[1], S. Galvani[1,2,*], P. Alemi[1]

[1]*Department of Power Engineering, Urmia Branch, Islamic Azad University, Urmia, Iran.*

[2]*Department of Power Engineering, Faculty of Electrical and Computer Engineering, Urmia University, Urmia, Iran.*

Article Info

Article History:
Received 18 September 2019
Reviewed 11 November 2019
Revised 02 January 2020
Accepted 04 April 2020

Keywords:
Economic Dispatch
Energy Hub (EH)
Configuration
Demand response
Metaheuristic Automatic Data Clustering (MADC)

*Corresponding Author's Email Address:
s.galvani@urmia.ac.ir

Abstract

Background and Objectives: Different energy demand calls the need for utilizing Energy Hub Systems (EHS), but the economic dispatch issue has become complicated due to uncertainty in demand. So, scenario generation and reduction techniques are used to considering the uncertainty of the EH demand. Dependent on the amount of fuel used, each system has various generation costs. Configuration selection stands as a challenging dilemma in the EHS designing besides economic problems. In this paper, the optimal EHS operation along with configuration issue is tackled.

Methods: To do so, two EHS types are investigated to evaluate the configuration effect besides energy prices simultaneously change. Typically, the effect of the Demand Response (DR) feature is rarely considered in EHSs management which is considered in this paper. Also, Metaheuristic Automatic Data Clustering (MADC) is used to reduce the decision-making problem dimension instead of using human decision-makers in the subject of cluster center numbers and considering uncertainty. The "Shannon's Entropy" and the "TOPSIS" methods are also used in the decision-making. The study is carried out in MATLAB© and GAMS©.

Results: In addition to minimizing the computational burden, the proposed EHS not only serves an enhancement in benefit by reducing the cost but also provides a semi-flat load curve in peak period by employing the Emergency Demand Response Program (EDRP) and Time of Use (TOU).

Conclusion: The results show that significant computational burden reduction is possible in the field of demand data by using the automatic clustering method without human interference. In addition to the proposed configuration's results betterment, the approach demonstrated EH's configuration effect could consider as important as other features in the presence of DRPs for reaching the desires of EHs customers which is rarely considered. Also, "Shannon's Entropy" and the "TOPSIS" methods integration could select the best DRP scenario without human interference. The results of this study are encouraging and warrant further analysis and researches.

Introduction

A. Motivation

The EHSs or briefly energy hubs (EHs) [1] could consider as a form of integrated distributed generations (DGs) [2], [3] which meet a variety of demands. Imported natural gas and electrical energy are typically the major supplies to these devices. EH as the system for managing power

grids is most important for its role in future networks, energy management systems, and Demand Response Programs (DRP) [4] along with achieving economic goals. As one of the most modern developments in the power systems, EHs have been commonly used in numerous implementations with diverse purposes to satisfy the needs of different demands like cool, heating, and electricity together.

To satisfy the various types of loads listed above, the EHs may be used, including various types of energy services such as upstream grid, Combined Heat and Power (CHP) units, boilers, and others. It should be mentioned that the requirements of energy systems can be different. While economic concerns have always been the first concern in the scheduling and management of power systems problems, a limited deal of effort has recently been made to recognize the role of EHs configuration effect in the optimal management and economic benefits EHs besides considering DRPs [5].

EH structures may include industrial plants, large housing complexes, or rural and urban districts. From an operating point of view, a key problem is the efficient management of such an organization (e.g., prices, peak reduction, and other elements).

The efficiency of the hub solution offers considerable management opportunities. For example, it is feasible, to stop using the particular equipment because of competitive electrical energy prices from the power grid at some special hours. The EH tends to be dynamic and competitive in terms of demand sensitivity. This issue could be a beneficial attribute for the introduction of EHs [6]. Concerning the mentioned questions, optimal operation of EH considering different structures was developed, where the concentration is on configuration effects to objective function in the DRPs presence and applying MADC. The EH configuration's effect [7], [8] rarely evaluated as an effective way of optimizing the mentioned objectives alongside other opportunities especially in presence of DRPs. As well, the DRP needs an approach for classifying the big data of demands which here is automatically clustering. This is while the system needs a pre-decision plan to participate (or not) in DRP. The reliable results will be encouraging to use the system of automatic clustering on conventional platforms to reduce the costs of analyzing big data. In this vision, it is important to find out the data cluster centers automatically for using in DRP decision-making, instead of human decision-making methods which in some of the research considered.

B. Literature Survey

Previously, EHs have been researched and their reviews are briefly described as follows:

Some papers like [9] analyzed price and security balance in power markets and proofed that this balance depends on prices in involuntary load shedding mode. It is said that by increasing in price the customer less notice to security and reliability of the system. The optimal economic operation of the EH has been determined in [10]. In the paper [11], a novel matrix modeling approach was suggested to promote the computerized simulation of multi-energy structures. In [12], the authors used a heuristic-based optimization algorithm called time variable acceleration coefficient-gravitational search algorithm to solve the power flow problem of the EH. To minimize the overall cost of EH operation, a robust optimization approach is employed in [13]. The paper [14] is about gas transmission [15] but so close to subject analysis the successive linear programming (SLP) for economic load dispatch. The paper [16] discussed gas and electricity mutual effect. This influence directly affects system security as paper [[17] evaluated.

The consequence of gas price increase is an increase in economic dispatch prices. This is reasonable because the gas price depends on fossil fuels. Likewise, fossil fuels are affected by electricity prices in energy markets.

This is because of that the most usage of gas is by power plants. By considering most of the works published about EH, the most popular inputs considered are electricity and gas, also it is proposed to use especially renewable energy in future investigations is out of the scope of this paper.

Paper [18] used a special model for EH economic dispatch. In this paper, some CHPs and other system parts like the furnace, transformers, compressors, and heater exchangers modeled.

The paper [19] compounds EH and DR and simultaneously try to use this concept by considering load shedding and other roles in energy management. For achieving the best result, this paper considered weather data, load data log, and fuel curves. Fig. 1, shows the mentioned paper's concept by using Supervisory Control and Data Acquisition (SCADA) center. This figure could be mentioned as the main idea of EH investigations. Paper [20] used different strategies for analyzing the compound share of wind, gas, and electrical energy as an input in the new structures of energy systems [21]. The model of the EH model matrix deliberate as Fig. 2 [22]. This matrix is like other popular energy systems. The input is connected by an energy converting box to output. Further, this idea will be detailed. Typically, conventional networks are hierarchical. In this structure, the input and output never interact. The main ring, which connects the future vision of the network to the conventional form, are the parts like CHP and Renewable Energy Sources (RES). This kind of network faces some problems like low power quality, complexity, protection problems, and environmental concerns [23]. The structures of new networks are nonhierarchical [24].

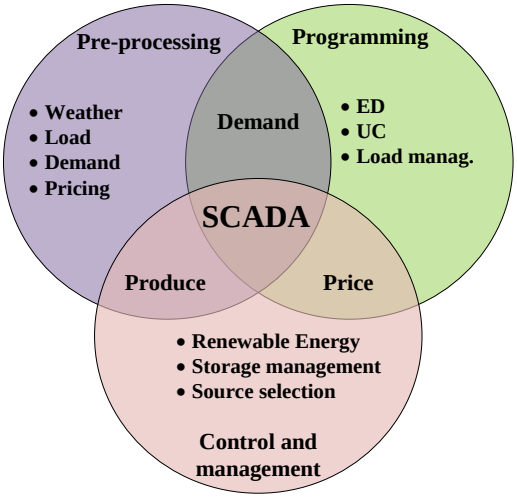


Fig. 1: Integrated energy management system.

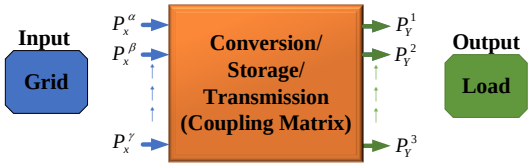


Fig. 2: Energy hub model.

The future structure of power systems is depicted in Fig. 3. In this figure, EHs are the interface between participation and transmission systems. So, EH and related forms which are bases of it—like energy management systems, CHP, DRP, etc.—are the forecasted system for future networks. Fig. 4 depicts a conventional model of industrial EH which could be an industrial site.

Demand-side is also an important subject for numerous researchers. A variety of solutions can be used as a supply-side uncertainties management in power systems. Since the short-term and long-term planning uncertainties of EH demand is not deniable, power system decision-makers and operators have used different uncertainty handling strategies, as described by [25]. The key difference between these instruments is related to the various approaches used to characterize the uncertain parameters. Stochastic models, for example, use the Probability Density Function (PDF) to model an uncertain parameter, while the fuzzy approach uses membership functions to define it [26]. In some references, the Monte Carlo approach was used to obtain the best precision [27]. Some researches, on the other hand, have focused on other approaches to finding effective solutions, including the point estimation method [28]. In [29], the point estimating approach is used to tackle the stochastic nature of renewable generating systems, and the demand uncertainty is provided by a robust optimization approach. However, compared to the above approaches, a variety of experiments have used a scenario-based approach to achieve acceptable accuracy [30], [31], and [32].

However, the Monte Carlo Simulation (MCS) method can be used effectively for probabilistic evaluation, but it requires a huge computational burden, making it unsuitable for problems with online optimization in particular. The alternative techniques that present an acceptable level of accuracy are quick and easy to apply. Some of these alternative methods are the point estimation method, the method of data clustering [33] and the method of Latin Hypercube Sampling (LHS) [34]. The proposed approach in this paper (MADC) needs no specific knowledge of the data to be categorized, as opposed to most of the mentioned methods. Instead, it evaluates the optimum number of scenarios of the results which named cluster centers.

Economic dispatch modeling is used for economic trading between the cost of production and the cost of versatility to reach the highest degree of network efficiency in the presence of storage and related technologies as a kind of energy system. In the smart grid systems, Demand Response Resources (DRRs) are implemented as a virtual power plant to improve the adequacy of the power network. DRRs frequently struggle to reduce their load. In [35], the reliability model of the DRR is developed as a multi-state traditional generation unit, where the probability, frequency of occurrence, and departure rate of each state can be obtained. Using the principle of power to gas has been analyzed to reach economic objectives in [36]. To reduce the running costs of the microgrid-based energy center network, a real-time pricing method has been used by [37]. The problem of EH economic dispatch is discussed in [38].

C. Contribution and Novelty

The key objective of this study is to enhance the economic operation of the EH and to resolve the problem of DRP alongside suggesting an optimal configuration. Operational costs and EH configuration are interconnected. A thorough approach has therefore been developed to include a desirable solution for operational costs and DRP. Many techniques such as the K-means have been implemented for data clustering in previous papers. The K-means algorithm is one of the easiest and most common categorization algorithms. This approach is capable of classifying a vast amount of data and clustering is such that the overall size of each data to the closest center of the cluster is reduced [39]. Regardless of its advantages, the K-means cannot find the number of optimal clusters. By using the MADC approach, the proposed model is resolved and various answers are obtained. The mentioned solution is also a more unfailing method rather than the data clustering (which more depends on human attitude). In automatic data clustering, the demand side reduced scenarios will choose automatically by using metaheuristic algorithms.

Additionally, in this paper for more reliability, the

final solution is compared with base configuration outcomes in the presence of DRP to demonstrate the efficiency of the proposed configuration.

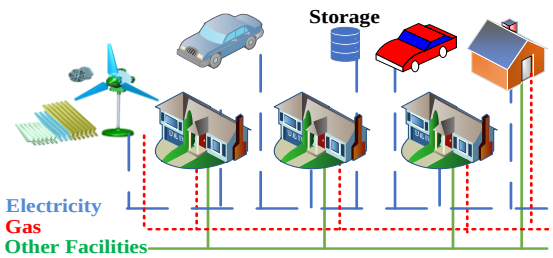


Fig. 3: Multi EHS diagram.

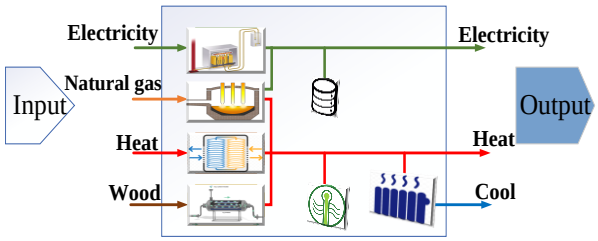


Fig.4: The Conventional Model of EHS.

As the Multiple Criteria Decision Making (MCDM) technique, the "Shannon's Entropy" and the "Technique for Order Preference by Similarity to Ideal Solution (TOPSIS)" methods are used to choose the best compromise solution from the solutions obtained. Briefly, analyzing the effect of energy hub configuration on DRP peak reduction considering MADC has been evaluated. The novelties of this paper are highlights as follows:

- Presenting the MADC through metaheuristic algorithms for managing the optimal operational performance of EHS in the presence of EDRP and TOU;
- Using automatic clustering instead of focusing on a large amount of data for considering uncertainty;
- Considering variable prices for electricity and gas simultaneously;
- Proposing an efficient configuration and analyzing the configuration effect in the presence of DRP to load shedding and other parameters;
- Investigating the benefits of the optimal configuration;
- Using "Shannon's Entropy" for weighting and the "TOPSIS" approach for the optimal scenario;
- Comparing the proposed configuration with the base configuration and encouraging and warrant further analysis and research.

D. Organization

The remainder of the paper is established as follows: The problem formulation and implementation which includes the associated constraints discussed in problem formulation section. Afterward, the system under study, including input data, assumptions, and the results of the EHS scheduling problem are presented. Likewise, this section studies the objective function. Eventually, the conclusion of the proposed EH and the discussion are presented and discussed in last section.

Problem Formulation and Implementation

A. Model of Conventional EH

As mentioned, modeling the EH is so similar to other systems. The model consists of three parts as all regular energy systems. The first part is input, then the process part, and finally the output block. The matrix in (1) will introduce this system as an integrated part. In this matrix, *C* is the coupling matrix, and *"α, β,* etc." represent energy carriers, and "*1, 2,* etc." represent various outputs.

The P_Y present output and the P_X on the other side is the input matrix which the dimension depends on the configuration.

$$\begin{bmatrix} P_Y^1 \\ P_Y^2 \\ \vdots \\ P_Y^n \end{bmatrix} = \begin{bmatrix} C_{\alpha\alpha} & C_{\beta\alpha} & \cdots & C_{\gamma\alpha} \\ C_{\alpha\beta} & C_{\beta\beta} & \cdots & C_{\gamma\beta} \\ \vdots & \vdots & \ddots & \vdots \\ C_{\alpha\gamma} & C_{\beta\gamma} & \cdots & C_{\gamma\gamma} \end{bmatrix} \begin{bmatrix} P_X^\alpha \\ P_X^\beta \\ \vdots \\ P_X^\gamma \end{bmatrix} \quad (1)$$

P_X as an input supplied by transformers and other subsystems of EH which is mentioned before. The middle matrix is the presenter of conversion, storage, and transmission part [22].

The thorough system can model by (2), which *P* is the input supplied by transformers and other systems, $\dot{E}$ is the storage part, *S* presents converters, and *L* leads to load. *C*, as introduced before, is the coupling matrix.

$$L = S.P - C.\dot{E} = \begin{bmatrix} S & -C \end{bmatrix} \begin{bmatrix} P \\ \dot{E} \end{bmatrix} \quad (2)$$

EH and its components, which consist of power resources, transmission, storage, and load management systems, are one by one a complete system. It is important to emphasize connections between smart grids main idea and EHs, where an EH could be a part of the smart grid. In the part of the converter, CHP has the most role. The EH concept could be a single house or an entire region of the city.

The formulas of each part of the system used, simplified, and linearized. The linearization can be focused on by whom favorite to accurate results.

- *CHP*

CHP is the most famous part of EH. Herein CHP receives the natural gas (G_t) and outputs heat (H_t) or electricity (E_t). This means

$$H_t = \eta_{gh}^{chp} G_t \quad (3)$$

$$E_t = \eta_{ge}^{chp} G_t \quad (4)$$

In the above equations η_{gh}^{chp} and η_{ge}^{chp} are the coefficient of heat and electricity in CHP.

- *Electric heat pump*

The Electric Heat Pump (EHP), gets electricity and gives heat (H_t) or cool (C_t) at one moment (not simultaneously). This means

$$C_t + H_t = E_t \times COP \quad (5)$$

$$H_t^{\min} I_t^h \le H_t^{EHP} \le H_t^{\max} I_t^{ch} \quad (6)$$

$$C_t^{\min} I_t^c \le C_t^{EHP} \le C_t^{\max} I_t^c \quad (7)$$

$$I_t^c + I_t^h \le 1 \quad (8)$$

$$I_t^c, I_t^h \varepsilon \{0,1\} \quad (9)$$

COP in the above equations is the coefficient factor, $H_t^{\max,\min}$ / $C_t^{\max,\min}$ are the low and high capacity of heat/cool generation of EHP.

- *Chiller boiler*

The Chiller Boiler (CB) receives heat and change to cool to provide cool demand. Here is the equation which η_{hc} is the coefficient of CB:

$$C_t = \eta_{cb} H_t \quad (10)$$

- *Electricity storage system*

The Electricity Storage System (ESS) is the most important part to provide flexibility in electricity provision, which is used as storage. ESS formulated as below:

$$SOC_t = SOC_{t-1} + (E_t^{ch} \eta_c - E_t^{dch} / \eta_d) \Delta_t \quad (11)$$

$$E_{\min}^{ch} \le E_t^{ch} \le E_{\max}^{ch} \quad (12)$$

$$E_{\min}^{dch} \le E_t^{dch} \le E_{\max}^{dch} \quad (13)$$

$$SOC_{\min} \le SOC_t \le SOC_{\max} \quad (14)$$

$$I_t^{dch} + I_t^{ch} \le 1 \quad (15)$$

$$I_t^{dch}, I_t^{ch} \varepsilon \{0,1\} \quad (16)$$

SOC_t and SOC_{t-1} are states of charge at moment *t* (which in this paper the dimension is 1 hour) and the moment before *t* which is *t-1*. $Soc_{max/min}$ is the high/low limit of these factors. $E_{\min/\max}^{ch,dch}$ and $E_t^{ch,dch}$ respectively are limitations of ESS charge or discharge at moment t, that means to get electricity from the network or give it to load, the $I_t^{ch,dch}$ control not to do this function at the same time. The charge and discharge efficiency showed by $\eta_{c/d}$. Notice that as the period is 1 hour, so 1 MWh=1 MW. ESS help to control the operation of the hub by charging/discharging in the necessary hours. Low price time is thus the correct time to charge, and high price time is used to avoid the purchase of electricity from the network. All electrical and gas resources are included in the operation of the hub and their optimum use was explored in order to reduce the running costs of the system.

- *Transformer*

Transformer (Tr) which receives electricity and give—in a different level of voltage—electricity, formulated as follows, where η_{ee} is the coefficient of Tr. It is important to notice that the change in voltage levels doesn't change in energy amounts. This means both sides -which are named the primary side and the secondary sides- are the same in energy amounts (except the energy loss), but the current and respectively the voltage change as mentioned (the current and voltage are out of scope in this paper). The $E_t^{in/out}$ represent the power of input/output.

$$E_t^{out} = \eta_{ee} E_t^{in} \quad (17)$$

- *Furnace*

Furnace (F) converts gas to heat by the coefficient of η_{gh}:

$$H_t = \eta_{gh} G_t \quad (18)$$

Mentioned parts summarized in the Table 1:

Table 1: Parameters of EHS equipment

Equipment	Output	Input
CHP	H_t^{CHP}	$\eta_{gh}^{chp} G_t^{in}$
	E_t^{CHP}	$\eta_{ge}^{chp} G_t^{in}$
EHP	$C_t^{EHP} + H_t^{EHP}$	$E_t^{in} \times COP$
Chiller boiler	C_t^{CB}	$\eta_{cb} H_t^{in}$
Transformer	E_t^{Tr}	$\eta_{ee} E_t^{in}$
Furnace	$H_t = H_{1,t} + H_{2,t}$	$\eta_{gh} G_t^{in}$

B. Case Study

The EHs including "base structure" and "proposed structure " are presented in two types as shown in Fig. 5 and Fig. 6. In type-A EHP is fed from the demand side but in type-B, EHP is fed from the input of EH. In other words, in the topology type-A, EHP has been fed as a part of total electricity demand which is presented by D_t^e for both types. Equation (19) shows the economic operation cost function which is introduced as the Objective Function (OF). The power balance equations for both types are presented as (20) and (21). Table 2 represents the variable of ED optimization variables.

C. Input Data and Assumptions

Daily demands and price [40] are as Table 3. The data used for both types. Demands per hour for heat, electrical and cool demands (MW) and carriers' price ($/MWh) change as shown by Fig. 7, which D_h is heating demand, D_e electrical, and D_c cooling demand.

Table 2: Variables of the optimization problem

Variable/Parameter	Description
E_1^t	ESS input for period *t*
E_2^t	Transformer input for period *t*
E_3^t	EHP input for period *t*
G_1^t	CHP input for period *t*
G_2^t	Furnace input for period *t*
λ_t^e	Electricity price for period *t*
λ_t^g	Gas price for period *t*

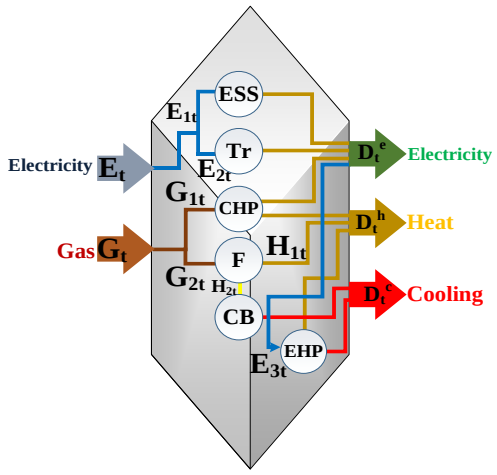


Fig. 5: EH configuration of type-A.

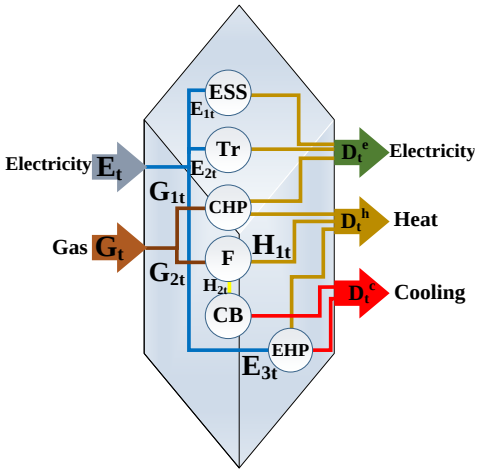


Fig. 6: EH configuration of type-B.

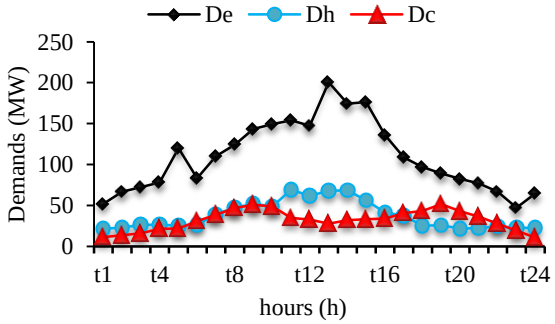


Fig. 7: Total electric, cool, and heat demand.

$$OF = \sum_t \lambda_t^e E_t + \lambda_t^g G_t \tag{19}$$

$$\begin{cases} if\ type\ A \rightarrow D_t^e = \eta_{ee} E_{2,t} + E_t^{dch} + \eta_{ge} G_{1,t} - E_{3,t} \\ if\ type\ B \rightarrow D_t^e = \eta_{ee} E_{2,t} + E_t^{dch} + \eta_{ge} G_{1,t} \end{cases} \tag{20}$$

$$\begin{cases} if\ type\ A \rightarrow E_t = E_{1,t} + E_{2,t} \\ if\ type\ B \rightarrow E_t = E_{1,t} + E_{2,t} + E_{3,t} \end{cases} \tag{21}$$

Table 3: Daily load demand and price

T (hours)	D_h (MW)	D_e (MW)	D_c (MW)	Electricity price $/MWh	Gas price $/MWh
t_1	21.41	52.10	11.51	22.02	5
t_2	23.21	66.70	13.68	24.24	5
t_3	26.09	72.20	16.01	23.1	6
t_4	26.72	78.37	21.42	22.8	6
t_5	25.59	120.20	21.97	24.12	6
t_6	26.45	83.48	30.80	23.16	7
t_7	39.54	110.40	38.94	31.38	7
t_8	47.28	124.29	46.78	40.38	8
t_9	52.12	143.61	50.97	42.3	8
t_{10}	49.13	149.28	48.86	39.72	11
t_{11}	69.26	154.19	34.77	43.98	11
t_{12}	61.97	147.30	32.68	36.48	11
t_{13}	68.04	200.71	27.77	37.92	14
t_{14}	68.56	174.37	32.02	42.48	15
t_{15}	56.40	176.54	33.22	37.86	15
t_{16}	41.32	136.11	34.13	31.5	15
t_{17}	37.43	108.71	40.78	34.2	16
t_{18}	25.44	96.90	43.56	29.52	16
t_{19}	25.66	89.08	51.48	28.5	16
t_{20}	21.94	82.49	43.15	29.7	16
t_{21}	22.44	76.93	36.49	31.86	16
t_{22}	24.63	66.85	27.68	30.96	18
t_{23}	22.72	47.17	19.14	30.3	20
t_{24}	22.59	64.67	11.04	21.84	20

D. Constraints

The base test system (Type-A) is chosen for the analysis of EH properties, based on [40]. Type-B is the proposed configuration. This structure will be chosen by looking at reducing cost function and other objectives.

The structure is based on the assumptions that follow:

Analysis of the system in a stable state.

- The power flow through the converters is described as just power and efficiency.
- Losses are evaluated only as efficiency factors for each element.
- The performance of the converter systems is assumed to be constant.
- Reverse control flow doesn't occur.
- The coupling matrix is not normally invertible

(underdetermined equation structure).

- In the household EH under analysis, the natural gas obtained at the input ports of the hub is divided into two paths, one path to the supply of CHP and the other path to the furnace.
- The gas-consuming components which are the CHP, EHP, and furnace unit provide the need for electricity.
- Demands and price of the energy carrier are obtained over a regular day of the market (unless stated otherwise, as in the clustering part)
- The price and demand for both systems change identically to achieve a sustainable position.
- Opposite of [40] which the price of the gas is constant, the gas price varies but the average is the same {12 $/MWH} as the base EH).

The architecture of the EH studied in this paper is linear. It should be assumed that the same outcome (with more accuracy) could be obtained by applying the proposed approach to a nonlinear problem. Also, results are encouraging and warrant further analysis and research. Since the nonlinear problem is beyond the scope of this article, and the emphasis is only on the demand side and MADC relating to the economic dispatch of DRP, in this work, after evaluating cluster centers by MADC, the problem will be tackled by using Mixed Integer Linear Programming (MILP), in GAMS© and using "CPLEX" [41] solver. The "CPLEX" Optimizer as its simplex method used in the "C programming language" is used. However today it still supports many forms of mathematical optimization and provides interfaces other than "C". The analysis is applied on a Windows 10© PC with a 2.6 GHz 7-core processor and 16 GB RAM. The analysis is carried out in GAMS© and MATLAB© to incorporate MADC, EH Economic Dispatch (EHED), and DRP. The average simulation time is 69.40 sec for MATLAB© and less than 10 seconds for GAMS©. The model suggested for household consumption may also be used for other different applications.

All of the data series assumed that is obtained by using a sampling cycle T_s equal to one hour, for an operational horizon which here is a typical single day. The electricity prices of energy depend on the hour of the day, with a "high" value of 43.98 $/MWh applied at all hours from 00:00 to 24:00 and a "low" value of 21.84 $/MWh. Gas costs as mentioned change from 5 to 20 $/MWh. It is important to notice that the average is 12 $/MWh which is like the mentioned reference [40].

Other constraints summarized as following tables which directly get from [40]:

Table 4: Efficiency data

$\eta_{ch/dch}$	η_{ee}	η_{ge}	η_{gh}	η^{f}_{gh}	η_{hc}	W_{ehp}
0.9	0.98	0.35	0.45	0.9	0.95	2.5

Table 5: Capacity data

Capacity	SOC_{min}	SOC_{max}	ESS initial energy	$E^{ch/dch}_{min}$	$E^{ch/dch}_{max}$
	120 MW h	600 MW h	120 MW h	0 MW	120 MW
Capacity	CHP	C/H^{ehp}_{min}	C/H^{ehp}_{max}	Furnace	I
	250 MW	0 MW	500 MW	500 MW	0/1 -

The demand curve is usually distributed. It can be concluded that the distribution of the average daily is usually normal [42]. The approach used in this paper to produce data clusters which will be discussed in the following sections. Additionally, as data are directly derived from [40], there are no details about the consumption for MADC implementation. In order to investigate the uncertainty, overall uncertainty modeled by MADC to scenarios reduction. As a consequence, the usual distribution used to generate data for which simulate the sampling cycle T_s mentioned before. The following formula shows the probability density function (PDF) of a conventional load [42]. Electrical heating and cooling loads are modeled using the typical PDF:

$$PDF(L)=\frac{1}{\sqrt{2\pi\sigma_L^2}}e^{\frac{(L-\mu_L)^2}{2\sigma_L^2}} \quad (22)$$

In the above equations σ_L and μ_L specify the standard deviation and mean, respectively, L expresses the load value as well. The mean stands "the mean of the demands in a particular period" which the data collector saved and sigma is assumed to be %5.

E. DRP Modeling

Demand Side Management (DSM) as one of the most significant techniques used to maximize the benefits of electricity market players. DSM is called DR in deregulated power systems. Programs are typically divided into one of two categories: Incentive-Based Programs (IBP) and Time-Based Programs (TBP). Time-based pricing systems consist of the following schemes and the price of energy varies over times [43]:

- Real-Time Pricing (RTP),
- Time of Use (TOU),
- Critical Peak Pricing (CPP).

Incentive-based programs include:

- Interruptible/Curtailable service (I/C),
- Capacity market Program (CAP),
- Direct Load Control (DLC),
- Demand Bidding (DB),
- Emergency Demand Response Program (EDRP),
- Ancillary Service (A/S) programs.

Two DR mechanisms were mainly focused in this paper: TOU and EDRP. Also, DR is modeled based on the principle of load elasticity, considering TOU and EDRP

approaches, respectively using the multi-period load models which will consider in the following section. The suggested model is based on the EH model and the optimal rates are calculated for the TOU system (with the variable price of gas and electricity) as well as the optimal benefits for the integrated TOU and EDRP schemes. In the EDRP scheme [43] based on historical demand data, price data, and short-term load forecasting, Independent System Operator (ISO) seeks to reduce peak demand. Large EHs that want to reduce their consumption based on ISO announcements will participate in these programs. The ISO will pay them a significant amount of money (sometimes 10 times the electricity price in the off-peak period) as an incentive. It is obvious that customers will participate in this program voluntarily. This will raise a great deal of uncertainty about the peak reduction, but due to the pre-determination of the incentive amounts and also because there is not any penalty price for consumers who do not reduce their consumption, participation in this program has been very good in most systems. The ISO was able to return the price to its normal value by forecasting the load curve for other days out of the working days of the DRP. As a result, peak loading and price reduction are the program results. Electrical consumers can participate in EDRP in the energy market, to reduce their costs. In these processes, customers attempt to move their demands of electricity from peak to off-peak. The current electrical charge is equal to the main load plus the variable load according to DRP. These factors may be a decrease or increase in load either positive or negative value. The amount of load increase or decrease that is the percentage of load participation in EDRP should be subject to a predefined limit. Simultaneous load increase/decrease is not allowed, however. According to the fundamental rule of EDRPs, the amount of the moving demand will be almost zero for a complete cycle of service which here is 24 hours. The DRP equations could summarize by the following equations:

$$DRP_t = \frac{P_t^D - P_t^{DR}}{P_t^D} = P_t^D - P_t^0 \quad (23)$$

$$-DRP^{\max} \times P_t^0 \le DRP_t \le DRP^{\max} \times P_t^0 \quad (24)$$

$$\sum_{t=1}^{T} DRP_t = 0 \quad (25)$$

In the TOU plan, energy prices are assessed based on the cost of production. Consequently, the price will generally be inexpensive during the low loading period, moderate during the off-peak period, and high during the peak period. By operating this scheme, consumers, who can move their consumption, will adjust to their prices. As a result, peak demand will be reduced and loads will shift from peak to off-peak or low periods [43].

As the consumers have been separated from the effects and market behavior, "elasticity" as the determiner of the customers' behavior is characterized as a price-sensitive demand [43]:

$$Elasticity = \frac{\partial q}{\partial \rho} = \frac{\rho_0}{q_0} . \frac{dq}{dp} \quad (26)$$

Where q is the demand value (MWh), ρ is electricity energy price ($/MWh), ρ_0 is the initial electricity energy price ($/MWh) and q_0 presents the initial demand value (MWh).

When the price of electrical energy varies over various times, the market may respond to one of the following:

- Some loads are not able to switch from one time to another (e.g., lighting loads) and maybe "on" or "off" only. So, such loads are flat and it's called "Self-Elasticity," so it has always a negative value.

$$Self-Elasticity = \frac{\Delta D_j}{\Delta \rho_j} \le 0 \quad (27)$$

- Some consumption could be moved from high to off-peak or low times. This action is called Multi-Period Sensitivity and is determined by "Cross-Elasticity." This value is always positive.

$$Cross-Elasticity = \frac{\Delta D_j}{\Delta \rho_j} \ge 0 \quad (28)$$

where in the above equations ΔD_j is demand changes in period j, and $\Delta \rho_j$ represents price changes in period j.

The elasticity coefficients for hours of the day can then be represented in a 24×24 matrix by the Table 6 role which assumed as in [44]:

Table 6: Elasticities

	Peak	Off-peak	Low
Peak	-0.02	0.0032	0.0024
Off-peak	0.0032	-0.02	0.002
Low	0.0024	0.002	-0.02

The method of modeling and formulating how the DRP system impacts the market for energy and how the full gain to consumers is reached has been discussed [43], [44].

Also, details of the demand response economic model and the effect on electricity consumption, which is focused on optimizing the benefits are described in the mentioned papers.

The related sensitive economic final model of the load is thus presented as follows:

$$d(i) = \begin{Bmatrix} d_0(i) + \sum_{j=1}^{24} E_0(i,j) . \frac{d_0(i)}{\rho_0(j)} \times \\ A(j) + \frac{E(i)\left[\rho(i) - \rho_0(i) + A(i)\right]}{\rho_0(i)} \end{Bmatrix} \quad i = 1,2,...,24 \quad (29)$$

where $d_0(i)$ is demand in i-th hour (MWh), $\rho_0(i)$ presents electricity price in i-th hour ($/MWh), and $A(j)$ is the incentive in i-th hour ($/MWh). By considering the mentioned equations and definitions, the introduced EHs could model in the presence of the DR management system by Fig. 8. The detailed line connection in each type is presented in Fig. 9 and Fig. 10. The Mentioned equations indicate how high the customer's demand will be in order to achieve the full benefits within 24 hours. In the numerical results section, incentives could shift the demand curve when EDRP and TOU programs are running.

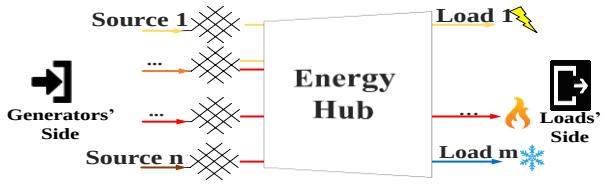


Fig. 8: Conceptual model of energy hub demand response.

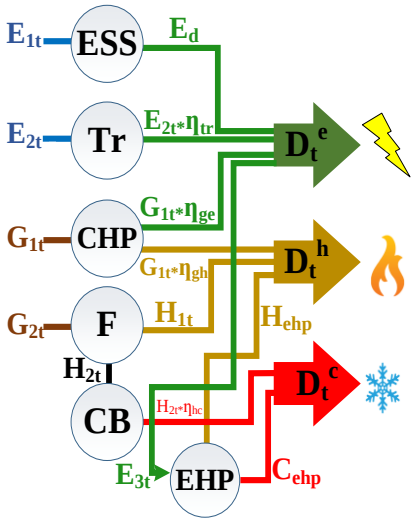


Fig. 9: Energy Hub type A's demand response scheme.

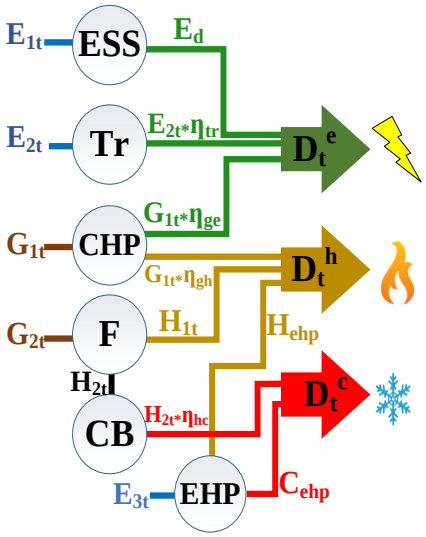


Fig. 10: Energy Hub type B's demand response scheme.

F. MADC

Clustering or unsupervised learning is the method of grouping data items into different partitions or clusters. In other terms, clustering identifies feasible cluster centers of multidimensional data based on some measure of uniqueness. This paper discusses the implementation of automated clustering [45] of large data sets which here is demand in the context of DRP. The proposed automatic clustering method does not require the preceding label of the data to be categorized. Also, it calculates the optimum number of data partitions automatically based on the metaheuristic algorithm laws. For clustering the data as mentioned in the previous section, the normal PDF applied to available data to produce demand scenario data. The most common approach to determine the similarities between two cluster centers is a distance measurement. The cluster validity indexes refer to the statistical-mathematical functions used to quantitatively test the performance of a clustering algorithm.

Table 7: MADC DB index approach outline

DB index	
Cluster scatter within i-th	$S_{i,q}=\left[\frac{1}{N_i}\sum_{\vec{X}\in C_i}\left\|\vec{X}-\vec{m}_i\right\|\right]^{\frac{1}{q}}$
Cluster distance between i-th and j-th	$d_{i,j,t}=\left[\sum_{p=1}^{d}\left\|\vec{m}_{i,p}-\vec{m}_{j,p}\right\|\right]^{\frac{1}{t}}$
$\vec{m}_i$ =i-th cluster center, N_i=the number of elements in the i-th cluster C_i, q,t>0, P=data points, $\vec{X}$ = data points as a matrix.	
R index	$R_{i,qt}=\max_{j\in K,j\neq i}\left\{\frac{S_{i,q}+S_{j,q}}{d_{ij,t}}\right\}$
DB index	$DB(K)=\frac{1}{K}\sum_{i=1}^{K}R_{i,qt}$ `
Fitness function	$F=\frac{1}{DB_i(K)+eps}$

The cluster validity index usually has two functions. First, it can be used to calculate the number of clusters, and second, it can be used to identify the best cluster centers. Two of the well-known indexes used in the literature for crisp clustering are the "DB" (Davies-Bouldin) index and the "CS" (Chou-Su and Lai) index [45]. Due to their optimizing nature, cluster validity indexes are better used in combination with any optimization algorithms such as GA, PSO, etc. In this paper, the DB index integrated with the GA algorithm has been used in the analysis of DRP for finding out the configuration

effect of EHS, because of its achievements in multi demand EHs [46]. Through using clustering instead of focusing on a large amount of data, only specified categories are evaluated. The comprehensive process of modeling and formulating the DB index has been discussed in [45], which can be used for more explanation. Table 7 summarized the mentioned method.

The MADC 's superiority is demonstrated by [45] comparing it with two established techniques of partitional clustering and one common hierarchical clustering algorithm. Besides, the objective function of DB and CS indices which introduced in this paper could present the optimal solution's guarantee based on the cluster scatter and distance eigen. Also, the key difference between these instruments is related to its ability to evaluate the optimum number of scenarios of the results which named cluster centers automatically without a human-deciding process.

"Shannon entropy" and "TOPSIS" Method Application

The Shannon entropy [47] can be used to assess the degree of disorder and its effectiveness in system information. The lower the entropy value, the lower the system's degree of disorder. The "Shannon entropy" weight approach is based on the amount of information needed to calculate the weight of the index and is one of the objective fixed weight methods. An entropy weight approach is used to evaluate the weight of the index in this paper, which is determined as follows. "TOPSIS" is the principle of identifying the optimal solution for decision-making problems, first of all, then finding a feasible and final solution and rating the solutions according to the similarity of the feasible solution to the optimal solution (positive or negative according to reduction or increasing need), finding the nearest solution to the ideal solution and the furthest from the negative one. The comprehensive process of modeling and formulating "Shannon's Entropy" and the "TOPSIS" methods have been discussed in [48], which can be used for more explanation. Table 8 and Table 9 abridged both approaches.

Results and Discussion

A. Optimal Configuration Selection Considering DRP

As mentioned in the structured EHS, EHP fed point distinct both types "A" and "B". Conditions of both case studies and optimal EH's operation cost has been presented through Fig. 11 to Fig. 17.

EHP converts electricity to heat and cool energy. Operation cost variation shows the effects of the new configuration in the EHP fed point. Results indicate that changing configuration is capable of reducing the operation cost to 0.06% in a single day. This reduction is achieved because of the EHP ability to manage energy. Operation cost changing demonstrates the fed point of equipment and consequently configuration influence on this subject. Only a little change in the fed point cause to reduce the hub's operation cost and it is obvious this reduction has been obtained with assumed demand and this percentage can change by differing mentioned conditions. Implementation of the selected scenario which is EDRP with $40 incentive rate in the following sections will result in %7.8-%9.9 cost reduction (operational cost for 1st cluster center is $74,925 and for 2nd one is $76,664 in the presence of DRP for type "B" in the selected scenario which will be discussed in the following section).

As shown in Fig. 11 and Fig. 13, the E_1 and E_3 inputs vary in the same way. This is because the EHP energy consumption is not related to the energy direction of the E_1 and E_3. On the other hand, the E_2 vice versa E_1 and E_3 changes in a different way for both types, which showed by Fig. 12 for both topologies. The E_1 constancy is because the E_1 is the basis of D_e energy consumption. The changelessness of E_3 is because of its role in feeding D_C, which in both types is the same.

It is necessary to compare the whole electricity energy usage in Fig. 14. It is important to mention again that the same demand is considered for both types of EHSs. Total input gas energy is depicted in Fig. 15. Heat demand production by the furnace and chiller boiler is illustrated in Fig. 16.

By utilizing the objective function which presented before, the energy hub economic dispatch will cost as shown by Fig. 17.

As can be seen, in the case without DR for both types (and just variable costs as different operational costs, not TOU applying), the value of operation cost for type "A" and "B" are $83,205.885 and $83,157.136, respectively. By comparing both cases, it can be found that, by employing type "B", operation cost reduced up to $48.74 which is 0.06%.

Table 8: Shannon entropy method outlines

Shannon entropy	
Normalizing matrix (X_{ij}= amount of eigen)	$r_{ij} = \frac{x_{ij}}{\sum_{j=1}^{m} x_{ij}}$
Entropy calculation	$e_{ij} = -\frac{1}{Ln(m)} \times \sum_{i=1}^{m} r_{ij} Ln\, r_{ij}$
Distance for full impact factor	$d_j = 1 - e_j$
Weight vector calculation	$w_j = \frac{1-e_j}{\sum_{j=1}^{n} 1-e_j}$

X_{ij}=each element of the decision matrix, m=the scenarios number

Table 9: TOPSIS approach outlines

TOPSIS	
Normalized specification matrix, $N=[n_{ij}]_{m\times n}$	$n_{ij}=\frac{x_{ij}}{\sqrt{\sum_{i=1}^{m}x_{ij}^2}}$
Weighted normalized specification matrix (weights from Shannon entropy)	$V=[w_j.n_{ij}]$
Identify the optimal value of each eigen where for maximizing proper V^+=max of each proper and V^-= min of each proper, and for minimizing proper V^+=min of each proper and V^-= max of each proper	$\begin{cases}V_j^+=[V_1^+,V_2^+,V_3^+,...]\\V_j^-=[V_1^-,V_2^-,V_3^-,...]\end{cases}$
Determine separation measures	$\begin{cases}S_i^+=\sqrt{\sum_{j=1}^{n}(V_{ij}-V_j^+)^2}\\S_i^-=\sqrt{\sum_{j=1}^{n}(V_{ij}-V_j^-)^2}\end{cases}$
Evaluate rank	$P_i=\frac{S_i^-}{S_i^++S_i^-}$
V_j^+ = optimal answer, V_j^- =worst answer	

By utilizing EDRP for peak hours (14-18) and different incentive rates, the mentioned types result as Fig. 18.

In type "B", the electrical energy input directions play a complementary role and help to cost reductions.

For comparing the results for both types, the simulation result of the optimization problem in deterministic conditions has been presented in Table 10, too.

These results imply that the hub's operation cost has a direct relation with the configuration.

The operational cost in type A is almost \$83,206. So, by employing type "B" and applying EDRP, operation cost reduced up to \$1244. In comparison with base type results (as Fig. 17), the operational cost for the least incentive rate (\$1) reduced 0.32% and for the highest incentive rate (\$20) reduced 1.50% in the selected type which is "B".

By considering the mentioned results type B is selected for applying the MADC approach. Also, the selected type is less complicated as presented in past sections. By considering both types as Fig. 19 presents, a configuration selection center could manage the effect of the configuration in reducing costs and emissions for a special site depending on situations (with DRP or without DRP and the combination of these configurations).

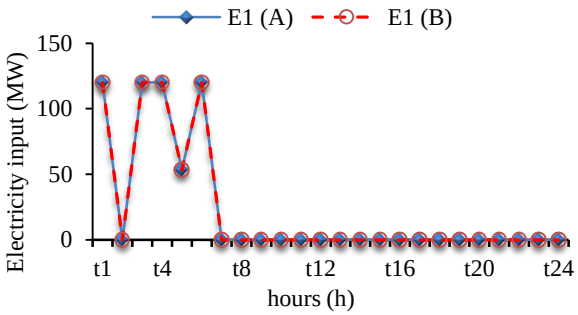


Fig. 11: E_1 consumption for both type-A and type-B.

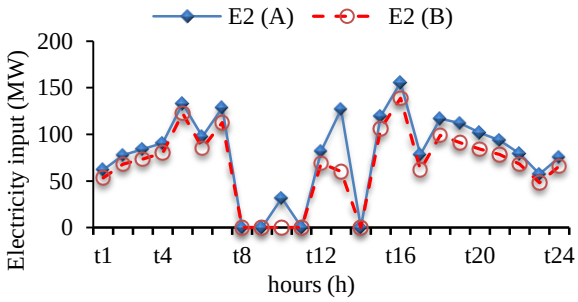


Fig. 12: E_2 consumption for both type-A and type-B.

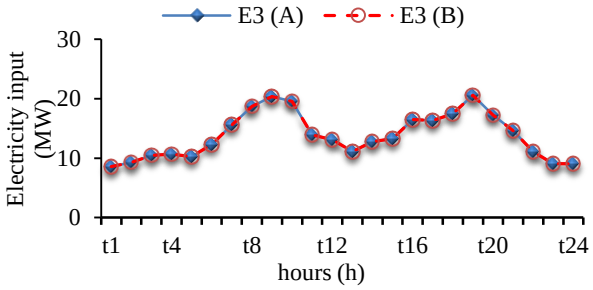


Fig. 13: E_3 consumption for both type-A and type-B.

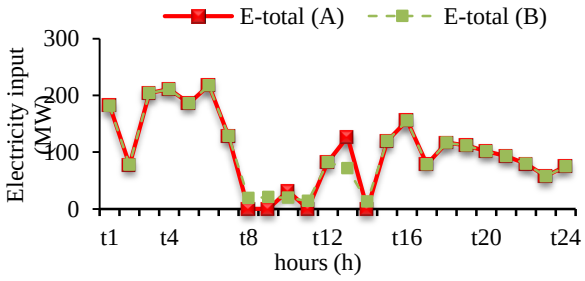


Fig. 14: Total energy input E for both type-A and type-B.

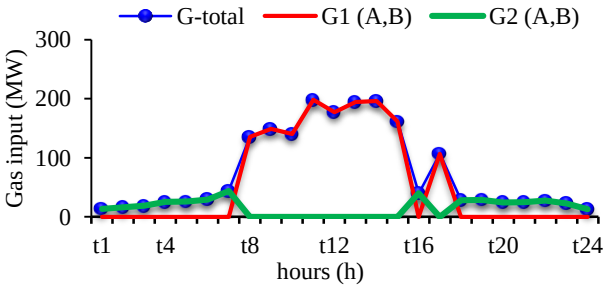


Fig. 15: Input gas energy for both types.

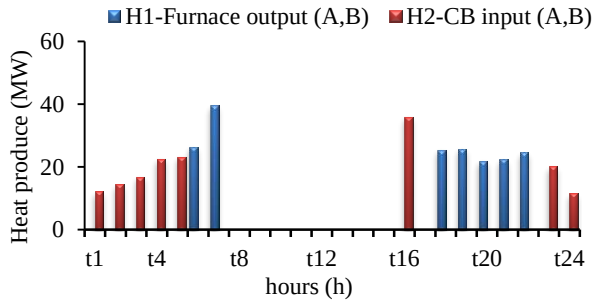


Fig. 16: Amount of produced heat by the furnace and chiller boiler for both types.

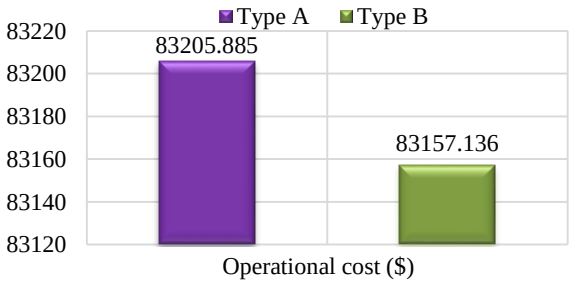


Fig. 17: Total operation cost of the base and proposed EH.

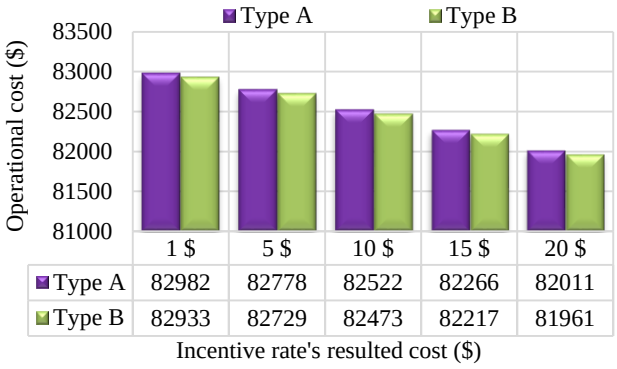


Fig. 18: Total operation cost comparison based on different incentives in the presence of EDRP.

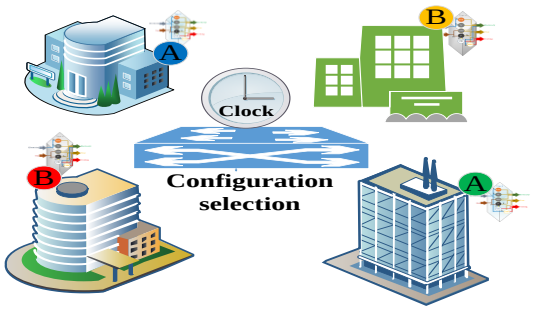


Fig. 19: Configuration selection scheme.

B. MADC Approach

In this section, MADC has been evaluated. In this paper for more reasonable results, the demands designate in a complete clustering act instead of clustering each hour [46]. For computing cluster center with MADC (by using GA_{DB} [46] as mentioned before) following assumptions used:

- Maximum number of iterations=200;
- Population size (nPop)=100;
- Crossover percentage (Pc)=0.8;
- Off springs number (parents)=2×round (Pc× nPop/2);
- Mutation percentage (Pm)=0.3;
- Number of mutants=round (Pm× nPop);
- gamma=0.05;
- Mutation rate=0.02;
- Selection pressure=8.

MADC was executed 100 times and the average run time is 69.40 sec, notice that the approach should run 3 times to result in cluster center combination of 3 demands (the last run results considered). Also, the average amount (costs) of the MADC fitness function (which is mentioned in Table 7) are 0.999, 0.990, and 1.014.

As it is illustrated in Fig. 20, by utilizing the automatic clustering method for each demand data, two cluster center results as each demand cluster centers. The illustrated data clusters as a complete view showed by Fig. 21. By considering the mentioned figures the final combination could be the as following matrix

$$Combinations = \begin{matrix} \textbf{combination 1} \\ \textbf{combination 2} \\ \textbf{combination 3} \\ \textbf{combination 4} \\ \textbf{combination 5} \\ \textbf{combination 6} \\ \textbf{combination 7} \\ \textbf{combination 8} \end{matrix} \begin{bmatrix} D_h Cluster\ 1 & D_e Cluster\ 1 & D_c Cluster\ 1 \\ D_h Cluster\ 2 & D_e Cluster\ 2 & D_c Cluster\ 2 \\ D_h Cluster\ 1 & D_e Cluster\ 2 & D_c Cluster\ 2 \\ D_h Cluster\ 2 & D_e Cluster\ 1 & D_c Cluster\ 1 \\ D_h Cluster\ 1 & D_e Cluster\ 1 & D_c Cluster\ 2 \\ D_h Cluster\ 2 & D_e Cluster\ 1 & D_c Cluster\ 2 \\ D_h Cluster\ 1 & D_e Cluster\ 2 & D_c Cluster\ 1 \\ D_h Cluster\ 2 & D_e Cluster\ 2 & D_c Cluster\ 1 \end{bmatrix} \quad (30)$$

For achieving the final results, the first cluster center combinations (1,2) are considered in this paper. Table 11 presents the selected cluster centers. Also, the results were analyzed by descriptive statistics in the mentioned table. For more reliability, the other combinations could consider that's beyond the scope of this paper. For future researches, the combinations could consider by two techniques. First, re-cluster detected clusters combination with an adaptable approach (like using metaheuristic algorithms), second using descriptive statistics features of evaluated cluster centers.

C. Different Scenarios in the Presence of EDRP and Prioritizing

Several DR scenarios have been considered for both evaluated cluster centers as indicated by Table 12 and Table 13. The suggested DRP is split into 6 scenarios. In the base case, base prices are implemented where no DR program is adopted as mentioned in the previous sections. It should be mentioned again that the price change in Table 3 is because of the operational cost of EH's equipment. Scenario #1 is the DRP without any variable price for electricity or gas (the average price in Table 3 considered for both which are $12 and $31.65 for gas and electricity respectively). Scenario #2 to

Scenario #4 are the IBP class, which includes the EDRP program with different incentive rates from $10 to $40. Finally, scenario #5 and Scenario #6 are the scenarios with %80 and %120 of elasticity and an incentive rate of $20.

In order to enhance the characteristics of the load profile as well as the customer's benefit, the following attributes are considered as elements 1 to 6 as seen in Table 12 and Table 13: "operational cost without DR ($), customer bill ($), operational cost reduction ($), costumer benefit ($), electrical energy peak reduction (MWh) and gas energy peak reduction (MWh)".

The attributes are weighted using the "Shannon entropy" method. The weights of the attributes measured are seen in Table 14. The decision matrix is then defined using "TOPSIS" with the results of Table 15. As mentioned, the decision matrix reflects the performance of each program for each attribute. As seen in Table 15, scenarios 4, 6, 3, 1, 5, and 2 give the best results respectively for both cluster centers. In order to avoid a vast number of statistics and tables from all the results of the scenarios, only results relating to the selected scenario (#4) have been presented and discussed in this section. The results of the simulation studies and the effect on the load curve characteristics of the selected DRP scenario (# 4) using the load economic model are shown in Fig. 22 and Fig. 23 for both cluster centers considering peak reduction in 14-18 periods. Load shifting -as one of the demand response strategies- has been successfully applied in all types of demand, electricity, heat, and cooling. The demand side action for both forms is acceptable. The detailed peak reduction for each demand is depicted in Fig. 24 and Fig. 25.

D. TOU and TOU+EDRP Integration Scheme for Cluster Centers

Typical EDRP has been completed with TOU in this paper for full utilization of the demand-side potential for decreasing operational costs. The TOU program is one of the most popular programs among the DRPs.

By using this program, the ISO can obtain optimum results with most benefits. Besides, the DRP control unit could adjust the energy usage from a particular time in the EH, which has a higher electricity price, to a night period with a lower electricity price with the use of demand response capability programs (which here this case considered by just doubling prices).

Most of the demands of the system are made up of external energy networks (so this could be a major possibility to change consume side behavior by adjusting energy price for gas/electricity).

Energy carriers have been described in this paper as electricity, natural gas.

As a consequence, this change impacts the power grid and the natural gas network. The initial price variability of energy carriers causes the EH management unit of the system to follow various strategies to achieve optimum benefit. As a special case for observing TOU effect Table 16 and Table 17 reflects the results of doubling just electricity price and as other case both electricity and gas prices for both cluster centers.

Fig. 26 to Fig. 29 are demonstrators of imported energy form the electrical /natural gas grid during the 24 hours for the selected type that is B. Applied approaches such as EDRP integrated with TOU changes the imported power amount significantly. In other words, the variation of electricity/gas price in proportion to its constant price manner results in providing the demand for EH in optimal scheduling.

Fig. 26 and Fig. 27 depict the twofold electricity price in peak for first and second cluster centers respectively. Also, Fig. 28 and Fig. 29 represent twofold electricity and gas price in peak for both cluster centers.

As seen doubling just electricity price (which in conventional systems could be even more) could reduce electricity usage to zero. In the other case by twofold electricity and gas price simultaneously the E_{in} and G_{in} reduced considerably.

E. Priorities for Future Researches

The impacts of other configurations could consider in future articles to get more certainty. Admittedly, by using more inputs including water and other facilities, it is possible to increase the baseline performance. Furthermore, the price profile (because of demand) may change on a seasonal, weekly, or even daily basis, so that the mathematical model may have to be adjusted almost daily depending on the operating conditions to avoid performance aberrations.

For each hour the uncertainties could be considered by greater reliability. Besides, the use of appliances such as CHP, turbine, etc.

to minimize the buying of energy from the power grid, is one of the solutions to increase the profit and reduce the expense of the energy center.

Analysis of the intermittent renewable energy production impacts could deliberate. Considering the natural energy, including wind turbines and PVs will improve the model. Thus, wind speed and solar radiation uncertainties will influence the outcome. The association and variability of the PV and wind turbines would have a direct effect on transmission lines and bus voltages.

Also, as mentioned the more cluster center combinations considering could increase reliability. So, for accurate, the forecast models the other available combinations could consider by re-cluster detected clusters combination or using descriptive statistics features of evaluated cluster centers.

The mentioned constraints could consider in further

works without changing the main mean and just by little changes in scenarios.

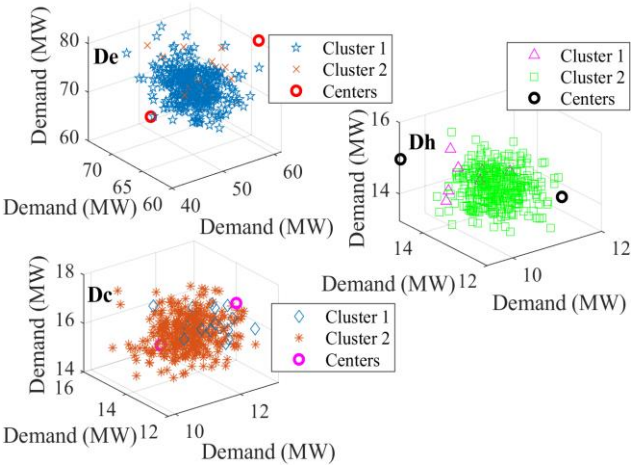


Fig. 20: Automatic clustering of demand's data.

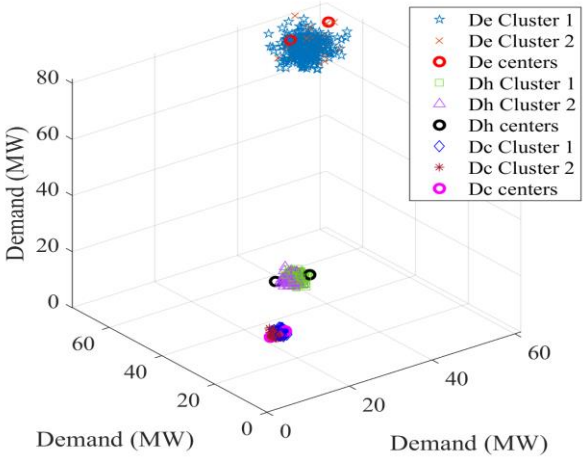


Fig. 21: Auto clustering of demands in a complete view.

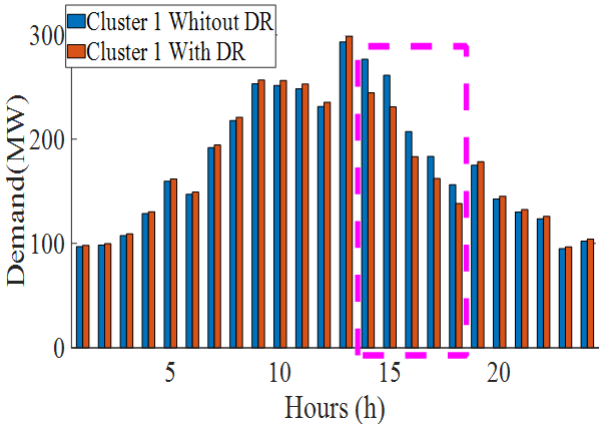


Fig. 22: EDRP peak reduction for 1st cluster center (#4 Type B).

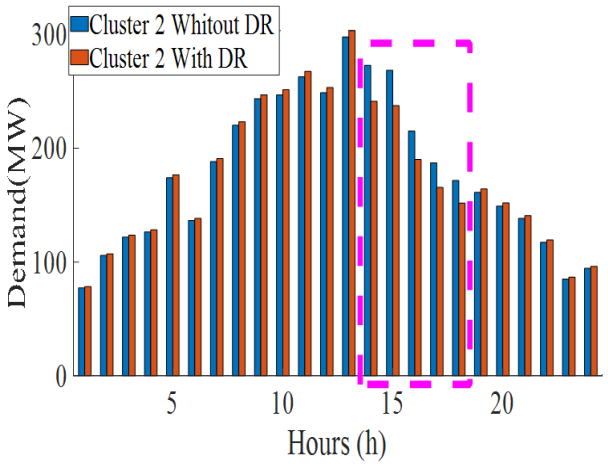


Fig. 23: EDRP peak reduction for 2nd cluster center (#4 Type B).

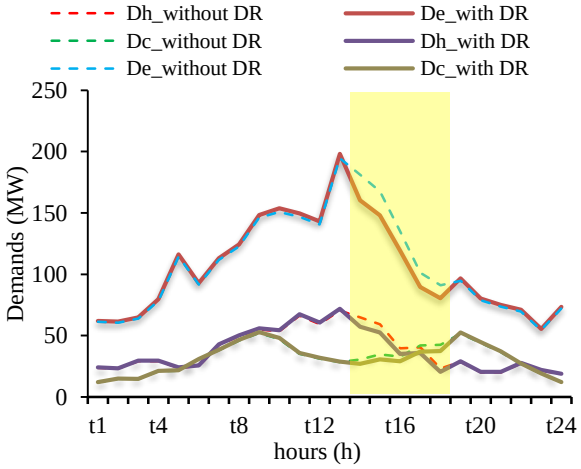


Fig. 24: Demands peak reduction 1st cluster center (#4 Type B).

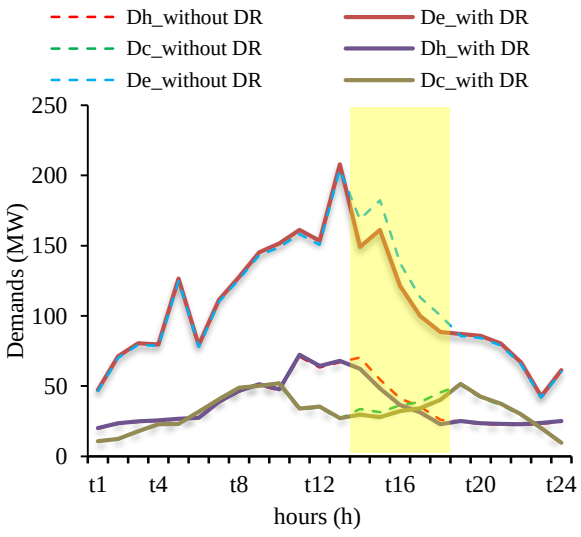


Fig. 25: Demands peak reduction 2nd cluster center (#4 Type B).

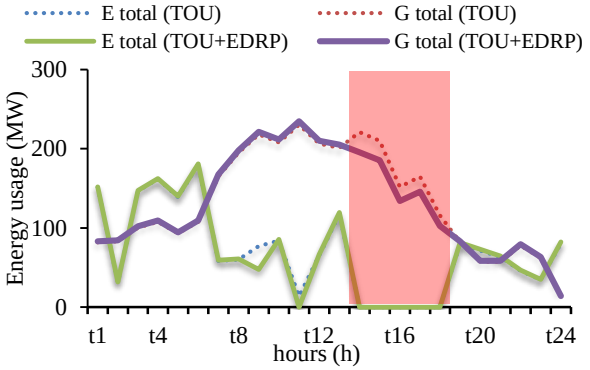


Fig. 26: Electricity and gas usage change 1st cluster center (twofold electricity price in peak).

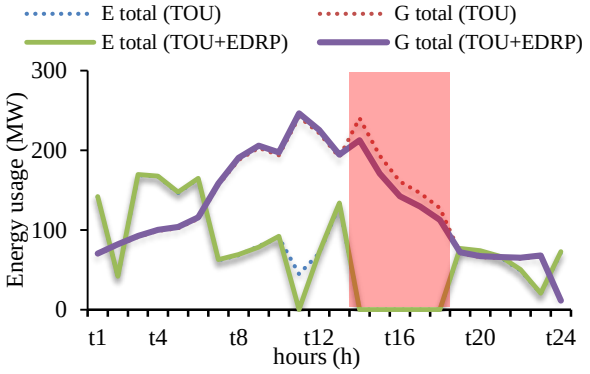


Fig. 27: Electricity and gas usage change 2nd cluster center (twofold electricity price in peak).

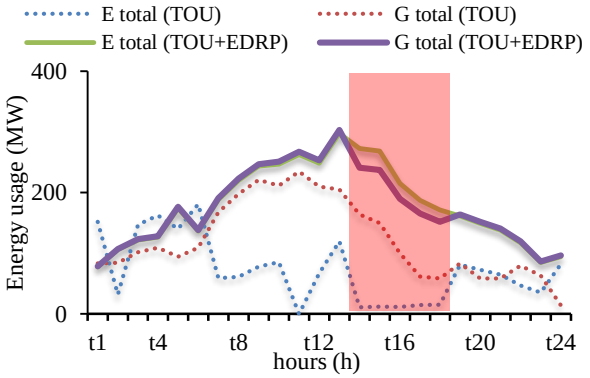


Fig. 28: Electricity and gas usage change 1st cluster center (twofold electricity and gas price in peak).

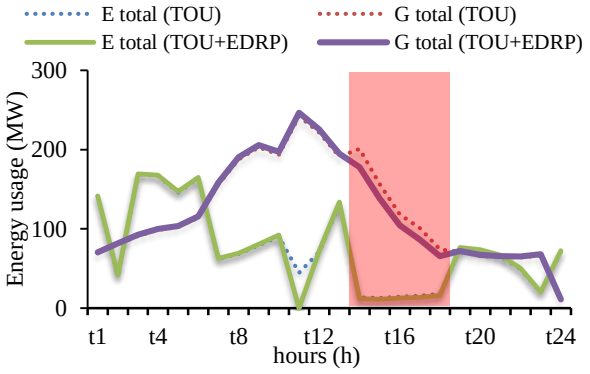


Fig. 29: Electricity and gas usage change 2nd cluster center (twofold electricity and gas price in peak).

Conclusion

In this paper demand-side uncertainties model of EH has been incorporated with selected DRP. The uncertainties of EH various demands including heating, cooling, and electricity have been defined as scenarios based on the PDFs of demands. Along with proposing DRP, the EH operator willing to minimize its costs has to determine the optimal scheduling of EHs as well as the selected DR scenarios based on two integrated approaches "TOPSIS" and "Shannon Entropy". The proposed problem has been solved by the MADC approach through the genetic algorithm and using the DB index. A typical EH has been employed to analyze the different aspects of the proposed method and improved by proposing some configuration changes. The effect of configurations considered in detail. It was shown that in the system which EHP was fed directly by the network, the total cost is less than the other structure. This is because of EHS acting by giving feedback from the output and reducing wasted energy and converting it to other kinds of energy. Also, the MADC model for EH in the presence of DR was developed to evaluate the prioritizing of DRPs and reducing the problem dimension. The "Shannon Entropy" method was used to obtain the weights. Then the "TOPSIS" method was used to select the best result and reasonably achievable by choosing available variables. Two cluster center combination models were presented and analyzed, with variables that met the EHED. The most important variables to evaluate the DRP of EH in the study area were the operational cost without DR ($), the customer bill ($), the operational decreased cost ($), the customer benefit ($), the peak reduced for electrical energy (MWh), and the peak reduced for gas energy (MWh). The costumer's benefit ($) is the most influential variable of MADM techniques for its weight that found by Shannon entropy for both clusters with a little tolerance.

Also, the "operational cost without DR ($)" was evaluated as the lowest weight in the decision matrix weight for both cluster centers. Besides, both cluster center as their origin -which is the normal PDF for demand as mentioned- act in weights almost the same. Although other variables are very important and should consider depending on the customer, ISO, or utility point of view, which here the customer considered. The user of this forecast model should use all of the available combinations in the clustering section since for more reliability the other combinations could consider by two techniques. First, re-cluster detected clusters combination with an adaptable approach, second using descriptive statistics features of evaluated cluster centers. Additionally, it is important to mention that when using the proposed models, it is necessary to consider seasonal, or weekend patterns which is out the

scope of this paper and could easily consider. So that the mentioned condition may have to be adapted almost daily depending on the conditions to avoid a performance defect.

The positive effect of employing DRP in both peak shedding and reducing economic costs presented in both types. By comparing the results for, economic cost, in the case with EDRP, is reduced in comparison to base type A. Besides, type B in the case with DRP is better in the cost function than type A in both with and without EDRP.

The inferences made of the study area showed that the EHED problem could consider by using cluster center instead of using large data in the presence of DRPs.

By employing the MADC and using DRP of EH the impact on the results of energy dispatch of EHs as well as the operating objective function have been determined and discussed. Besides, the scheduling results of resources after the realization of the most probable scenario have been illustrated. In brief, the simulation results show:

- In the DRP selection approach, the incentive rate has the most role than other elements like elasticity.
- MADC and selected algorithm by considering proper index (GA_{DB}) provided reduced scenarios of uncertainties by evaluating the optimum number of scenarios which named cluster centers (automatically without human-deciding process).
- The scenario finding process concentration is at the number of scenarios that will reduce the computation burden and increase the accuracy of the model.
- The power used by gas-based and electrical-based devices are considerably reduced by using DRPs during peak period.
- Using DRP leads to the reduction of the objective function alongside using optimal configuration. It means that motivating consumers to participate in DRPS can reduce the cost of EH operation.

Finally, this study can be a basis of comparison for future research in the area of study. As well, it is possible to use different mathematical models such as nonlinear and compare the response variable or the predictors without losing main achievements.

Author Contributions

H. Hosseinnejad, *S. Galvani*, and *P. Alemi* designed, carried out the data analysis, interpreted the results, and wrote the manuscript together.

Acknowledgment

The authors would like to thank Dept. of Electrical Engineering, Urmia Branch, Islamic Azad University.

Conflict of Interest

The author declares that there is no conflict of interests regarding the publication of this manuscript. Also, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancy have been completely observed by the authors.

Abbreviations

$\lambda_t^{e/g}$	Electric/Gas energy cost
η_{ee}	Coefficient related to transformer
$E_t^{ch/dch}$	Charging/Discharging in time *t*
$\eta_{ge/h/CB}$	Coefficient of gas to electricity/heat/CB
$H_{1/2,t}$	Furnace output for feeding heat load/CB
G_t	Input gas energy of EH in time *t*
$G_{1/2,t}$	Input gas energy directions
$G_{2,t}$	Input gas energy of the furnace
$D_t^{e/h/c}$	Electric/Heat/Cool demand at time *t*
E_t	Input electric energy in time *t*
$E_{1/2/3,t}$	Input electricity energy directions
$I_t^{ch/dch}$	Binary value to charging/discharge state
$I_t^{c/h}$	Binary value to cool/heat state
$C\ /\ H_t^{EHP}$	Cool/Heat generated by EHP
$COP = W_{ehp}$	The working factor of EHP
t	Sample time index
$P_{X/Y}$	Input/output power matrix
C	Coupling matrix
$DRP^{\max}$	Maximum participation limitation in DRP
$P_t^{D/0}$	Demand in time *t*
P_t^{DRP}	The available power of DRP
EHS	Energy Hub Systems
DR	Demand Response
MADC	Metaheuristic Automatic Data Clustering
EDRP	Emergency Demand Response Program
TOU	Time of Use
CHP	Combined Heat and Power
PDF	Probability Density Function
MCS	Monte Carlo Simulation
LHS	Latin Hypercube Sampling
DRR	Demand Response Resources
EHP	Electric Heat Pump
CB	Chiller Boiler
ESS	Electricity Storage System
Tr	Transformer
F	Furnace

Table 10: The DRP operation for both types with different incentive prices

	EH type	A	B	A	B	A	B	A	B	A	B
	Incentive price ($)	1.00		5.00		10.00		15.00		20.00	
Economical	Operational cost without DR ($)	33206.00	33157.00	33206.00	33157.00	33206.00	33157.00	33206.00	33157.00	33206.00	33157.00
	Operational cost with DR ($)	32982.00	32933.00	32778.00	32729.00	32522.00	32473.00	32266.00	32217.00	32011.00	31961.00
	Customer bill ($)	32978.80	32929.80	32698.08	32649.08	32202.34	32153.34	31546.76	31497.76	30732.30	30682.30
	Total paid incentive ($)	3.20	3.20	79.92	79.92	319.66	319.66	719.24	719.24	1278.70	1278.70
	Operational decreased cost ($)	224.00	224.00	428.00	428.00	684.00	684.00	940.00	940.00	1195.00	1196.00
	Costumer benefit ($)	227.20	227.20	507.92	507.92	1003.66	1003.66	1659.24	1659.24	2473.70	2474.70
Operational	Total E usage no DR (MWh)	1778.30	1775.00	1778.30	1775.00	1778.30	1775.00	1778.30	1775.00	1778.30	1775.00
	Total E usage with DR (MWh)	1772.635	1769.274	1769.082	1765.733	1764.64	1761.313	1760.196	1756.886	1755.757	1752.463
	Total reduced E (MWh)	5.68	5.68	9.24	9.22	13.68	13.64	18.12	18.07	22.56	22.49
	Total G usage without DR (MWh)	2967.00	2967.00	2967.00	2967.00	2967.00	2967.00	2967.00	2967.00	2967.00	2967.00
	Total G usage with DR (MWh)	2966.047	2966.047	2962.234	2962.234	2957.464	2957.464	2952.695	2952.695	2947.927	2947.927
	Total reduced G (MWh)	0.96	0.96	4.77	4.77	9.54	9.54	14.31	14.31	19.08	19.08
	Peak E usage without DR (MWh)	380.13	391.69	380.13	391.69	380.13	391.69	380.13	391.69	380.13	391.69
	Peak E usage with DR (MWh)	377.70	389.23	373.31	384.71	367.83	379.06	362.34	373.40	356.86	367.75
	Peak reduced E (MWh)	2.42	2.46	6.81	6.98	12.30	12.63	17.78	18.29	23.27	23.94
	Peak G usage without DR (MWh)	654.57	654.57	654.57	654.57	654.57	654.57	654.57	654.57	654.57	654.57
	Peak G usage with DR (MWh)	652.68	652.68	645.09	645.09	635.61	635.61	626.14	626.14	616.66	616.66
	Peak reduced G (MWh)	1.90	1.90	9.48	9.48	18.96	18.96	28.44	28.44	37.92	37.92

Table 11: Final selected cluster centers

T	1st cluster center			2nd cluster center			D_h probabilistic features				D_e probabilistic features				D_c probabilistic features			
(hours)	Dh (MW)	De (MW)	Dc (MW)	Dh (MW)	De (MW)	Dc (MW)	Mean	Std.Err	Std.Dev	Vari.	Mean	Std.Err	Std.Dev	Vari.	Mean	Std.Err	Std.Dev	Vari.
t1	23.7403	61.1934	12.0094	19.9127	46.5628	10.7536	21.83	1.35	1.91	3.66	53.88	5.17	7.32	53.51	11.38	0.44	0.63	0.39
t2	23.0667	60.5554	14.8005	23.2298	70.2743	12.2131	23.15	0.06	0.08	0.01	65.41	3.44	4.86	23.61	13.51	0.91	1.29	1.67
t3	29.2458	63.9393	14.4938	24.6736	79.5224	17.7252	26.96	1.62	2.29	5.23	71.73	5.51	7.79	60.71	16.11	1.14	1.62	2.61
t4	29.1081	78.4733	21.0230	25.3252	78.5930	22.5588	27.22	1.34	1.89	3.58	78.53	0.04	0.06	0.00	21.79	0.54	0.77	0.59
t5	23.7909	114.5180	21.4103	26.4597	124.9144	22.7427	25.13	0.94	1.33	1.78	119.72	3.68	5.20	27.02	22.08	0.47	0.67	0.44
t6	25.2912	91.5484	30.3805	27.0332	77.9971	31.4265	26.16	0.62	0.87	0.76	84.77	4.79	6.78	45.91	30.90	0.37	0.52	0.27
t7	42.2813	111.5222	38.0266	38.2310	109.8509	40.1385	40.26	1.43	2.03	4.10	110.69	0.59	0.84	0.70	39.08	0.75	1.06	1.12
t8	49.4742	122.5844	45.9311	45.9051	125.9513	48.2725	47.69	1.26	1.78	3.18	124.27	1.19	1.68	2.83	47.10	0.83	1.17	1.37
t9	55.0574	146.1714	52.0197	50.7099	143.1384	49.6187	52.88	1.54	2.17	4.73	144.65	1.07	1.52	2.30	50.82	0.85	1.20	1.44
t10	53.3959	150.8836	47.2973	46.8306	148.8455	51.0791	50.11	2.32	3.28	10.78	149.86	0.72	1.02	1.04	49.19	1.34	1.89	3.58
t11	66.1378	147.0003	35.2708	70.9739	158.2464	33.4281	68.56	1.71	2.42	5.85	152.62	3.98	5.62	31.62	34.35	0.65	0.92	0.85
t12	59.3973	140.5613	31.2797	63.2679	150.6599	34.7454	61.33	1.37	1.94	3.75	145.61	3.57	5.05	25.50	33.01	1.23	1.73	3.00
t13	70.4422	194.5400	28.3029	66.8692	204.1284	26.7481	68.66	1.26	1.79	3.19	199.33	3.39	4.79	22.98	27.53	0.55	0.78	0.60
t14	64.8269	181.3103	30.5516	70.3487	168.7717	33.5958	67.59	1.95	2.76	7.62	175.04	4.43	6.27	39.30	32.07	1.08	1.52	2.32
t15	59.2759	167.4717	34.5643	54.7006	182.2651	31.4916	56.99	1.62	2.29	5.23	174.87	5.23	7.40	54.71	33.03	1.09	1.54	2.36
t16	39.7297	134.8847	32.6926	41.3220	137.2725	36.3781	40.53	0.56	0.80	0.63	136.08	0.84	1.19	1.43	34.54	1.30	1.84	3.40
t17	40.5114	101.2534	41.8246	35.3257	112.9727	38.8061	37.92	1.83	2.59	6.72	107.11	4.14	5.86	34.34	40.32	1.07	1.51	2.28
t18	23.1794	90.9065	42.3504	25.9229	100.2231	45.4505	24.55	0.97	1.37	1.88	95.56	3.29	4.66	21.70	43.90	1.10	1.55	2.40
t19	28.6135	94.8067	51.6883	24.7605	85.6357	50.7225	26.69	1.36	1.93	3.71	90.22	3.24	4.59	21.03	51.21	0.34	0.48	0.23
t20	20.0601	78.6907	44.0028	23.0952	84.2897	41.6982	21.58	1.07	1.52	2.30	81.49	1.98	2.80	7.84	42.85	0.81	1.15	1.33
t21	20.0697	73.6350	36.5549	22.5983	78.8553	36.7498	21.33	0.89	1.26	1.60	76.25	1.85	2.61	6.81	36.65	0.07	0.10	0.01
t22	27.2648	69.7386	26.7349	22.4152	65.5656	29.2986	24.84	1.71	2.42	5.88	67.65	1.48	2.09	4.35	28.02	0.91	1.28	1.64
t23	21.7120	54.3546	18.8858	23.3281	41.9632	19.7832	22.52	0.57	0.81	0.65	48.16	4.38	6.20	38.39	19.33	0.32	0.45	0.20
t24	18.5066	71.9952	11.8792	24.6555	60.2725	9.5242	21.58	2.17	3.07	9.45	66.13	4.14	5.86	34.36	10.70	0.83	1.18	1.39

Table 12: The different scenarios of 1st cluster center for selected configuration (Type B)

1st cluster center	Scenario	Gas price ($/MWh)	Electricity price ($/MWh)	Incentive ($)	Elasticity	Operational cost without DR ($)	Customer bill ($)	Operational cost reduction ($)	Costumer benefit ($)	Electrical energy peak reduction (MWh)	Gas energy peak reduction (MWh)
	#01	12	31.65	20	×1	86390	85384	769.88	1775.80	26.87	37.66
	#02	Var.	Var.	10	×1	81451	81136	498.54	812.88	10.54	18.83
	#03	Var.	Var.	20	×1	80952	79695	997.03	2254.40	21.08	37.66
	#04	Var.	Var.	40	×1	79955	74925	1994.10	7023.60	42.17	75.31
	#05	Var.	Var.	20	×0.8	81151	80145	797.72	1803.60	16.87	30.13
	#06	Var.	Var.	20	×1.2	80753	79244	1196.50	2705.40	25.30	45.19

Table 13: The different scenarios of 2nd cluster center for selected configuration (Type B)

2nd cluster center	Scenario	Gas price ($/MWh)	Electricity price ($/MWh)	Incentive ($)	Elasticity	Operational cost without DR ($)	Customer bill ($)	Operational cost reduction ($)	Costumer benefit ($)	Electrical energy peak reduction (MWh)	Gas energy peak reduction (MWh)
	#01	12	31.6	20	×1	87707	86674	816.97	1850.20	28.47	37.67
	#02	Var.	Var.	10	×1	83392	83069	520.69	843.59	11.92	18.84
	#03	Var.	Var.	20	×1	82871	81580	1041.40	2333.00	23.84	37.67
	#04	Var.	Var.	40	×1	81830	76664	2082.90	7249.20	47.68	75.34
	#05	Var.	Var.	20	×0.8	83080	82046	833.15	1866.40	19.07	30.14
	#06	Var.	Var.	20	×1.2	82663	81113	1249.70	2799.50	28.61	45.21

Table 14: Weight of attributes matrix for selected configuration (Type B)

	Shannon entropy weight selection matrix of 1st cluster center						Shannon entropy weight selection matrix of 2nd cluster center					
	element 1	element 2	element 3	element 4	element 5	element 6	element 1	element 2	element 3	element 4	element 5	element 6
Selection matrix	-0.3058	-0.3070	-0.2578	-0.2409	-0.3143	-0.2879	-0.3049	-0.3061	-0.2597	-0.2418	-0.3075	-0.2879
	-0.2981	-0.3003	-0.2016	-0.1490	-0.1923	-0.1973	-0.2983	-0.3005	-0.2013	-0.1493	-0.1937	-0.1973
	-0.2972	-0.2979	-0.2927	-0.2729	-0.2824	-0.2879	-0.2974	-0.2981	-0.2924	-0.2730	-0.2840	-0.2879
	-0.2956	-0.2897	-0.3644	-0.3630	-0.3601	-0.3626	-0.2958	-0.2899	-0.3643	-0.3632	-0.3609	-0.3626
	-0.2976	-0.2987	-0.2626	-0.2429	-0.2522	-0.2578	-0.2978	-0.2989	-0.2623	-0.2430	-0.2538	-0.2578
	-0.2969	-0.2972	-0.3164	-0.2974	-0.3065	-0.3119	-0.2971	-0.2974	-0.3161	-0.2974	-0.3081	-0.3119
e_{j1}	0.9998	0.9995	0.9464	0.8742	0.9533	0.9519	0.9998	0.9996	0.94684	0.8750	0.9533	0.9519
$d_{j1}=1-e_{j1}$	0.0001	0.0004	0.0535	0.1257	0.0466	0.0480	0.0001	0.0003	0.05315	0.1249	0.0466	0.0480
W_{j1}	0.00067	0.00150	0.19506	0.45782	0.16996	0.17499	0.00052	0.00133	0.19453	0.45715	0.17058	0.17588
Improved weight	**0.0007**	**0.0015**	**0.1951**	**0.4578**	**0.1700**	**0.1750**	**0.0005**	**0.0013**	**0.1945**	**0.4571**	**0.1706**	**0.1759**

Table 15: Decision-making matrix of cluster centers for selected configuration (Type B)

Decision-making matrix of 1st cluster center					Decision-making matrix of 2nd cluster center				
S_{j1+}	S_{j1-}	P_{j1}	Rank	Scenario	S_{2+}	S_{j2-}	P_{j2}	Rank	Scenario
0.310661	0.077685	0.200040	4	#01	0.309590	0.076148	0.197409	4	#01
0.379267	0.000050	0.000132	6	#02	0.378545	0.000044	0.000116	6	#02
0.284599	0.096208	0.252643	3	#03	0.283844	0.096238	0.253203	3	#03
0.000000	0.379267	1.000000	1	#04	0.000000	0.378545	1.000000	1	#04
0.316010	0.063591	0.167520	5	#05	0.315329	0.063550	0.167732	5	#05
0.253568	0.129019	0.337228	2	#06	0.252761	0.129099	0.338078	2	#06

Table 16: TOU and TOU+EDRP scheme for 1st cluster center in selected configuration (Type B)

	1st cluster center (2×electricity price in peak)				1st cluster center (2×electricity price and 2× gas price in peak)			
	E total: TOU (MW)	G total: TOU (MW)	E total: TOU+EDRP (MW)	G total: TOU+EDRP (MW)	E total: TOU (MW)	G total: TOU (MW)	E total: TOU+EDRP (MW)	G total: TOU+EDRP (MW)
t1	151.296	81.876	151.731	83.014	151.296	81.876	151.731	83.014
t2	31.529	83.215	31.967	84.372	31.529	83.215	31.967	84.372
t3	146.875	100.511	147.249	101.909	146.875	100.511	147.249	101.909
t4	161.886	107.754	162.469	109.252	161.886	107.754	162.469	109.252
t5	138.976	93.015	140.167	94.308	138.976	93.015	140.167	94.308
t6	180.236	107.793	181.073	109.292	180.236	107.793	181.073	109.292
t7	58.327	165.279	59.138	167.577	58.327	165.279	59.138	167.577
t8	60.178	195.076	61.015	197.787	60.178	195.076	61.015	197.787
t9	76.922	218.148	47.385	221.181	76.922	218.148	77.991	221.181
t10	83.910	207.878	85.465	211.732	83.91	207.878	85.465	211.732
t11	13.533	230.218	0	234.485	13.533	230.218	0	234.485
t12	65.503	206.291	66.718	210.115	65.503	206.291	66.718	210.115
t13	117.414	201.263	119.591	204.994	117.414	201.263	119.591	204.994
t14	0	220.953	0	195.355	12.221	185.22	10.805	163.762
t15	0	209.786	0	185.482	13.826	169.36	12.224	149.739
t16	0	151.75	0	134.17	13.077	113.513	11.562	100.363
t17	0	164.665	0	145.588	16.73	115.747	14.792	61.604
t18	0	115.759	0	102.348	16.94	66.227	14.978	58.554
t19	79.877	81.753	81.358	83.268	79.877	81.753	81.358	83.268
t20	71.58	57.315	72.907	58.377	71.58	57.315	72.907	58.377
t21	63.429	57.342	64.605	58.405	63.429	57.342	64.605	58.405
t22	46.086	77.899	46.94	79.343	46.086	77.899	46.94	79.343
t23	34.533	62.034	35.173	63.184	34.533	62.034	35.173	63.184
t24	80.867	13.894	82.366	14.151	80.867	13.894	82.366	14.151
	Operational cost No DR		Operational cost with DR		Operational cost No DR		Operational cost with DR	
	69878.896 $		67893.945 $		78917.521 $		76397.990 $	

Table 17: TOU and TOU+EDRP scheme for 2nd cluster center in selected configuration (Type B)

	2nd cluster center (2×electricity price in peak)				2nd cluster center (2×electricity price and 2× gas price in peak)			
	E total: TOU (MW)	G total: TOU (MW)	E total: TOU+EDRP (MW)	G total: TOU+EDRP (MW)	E total: TOU (MW)	G total: TOU (MW)	E total: TOU+EDRP (MW)	G total: TOU+EDRP (MW)
t1	141.389	69.471	141.686	70.436	141.389	69.471	141.686	70.436
t2	41.232	80.655	41.805	81.776	41.232	80.655	41.805	81.776
t3	168.775	91.227	169.453	92.495	168.775	91.227	169.453	92.495
t4	166.971	98.742	167.624	100.115	166.971	98.742	167.624	100.115
t5	146.083	102.199	147.373	103.62	146.083	102.199	147.373	103.62
t6	164.123	113.994	164.736	115.579	164.123	113.994	164.736	115.579
t7	61.935	156.177	62.797	158.348	61.935	156.177	62.797	158.348
t8	68.296	187.616	69.246	190.225	68.296	187.616	69.246	190.225
t9	79.531	202.919	78.311	205.74	79.531	202.919	80.636	205.74
t10	90.444	193.543	92.12	197.131	90.444	193.543	92.12	197.131
t11	44.74	241.88	0	246.363	44.74	241.88	0	246.363
t12	70.73	221.403	72.041	225.507	70.73	221.403	72.041	225.507
t13	131.264	191.055	133.697	194.596	131.264	191.055	133.697	194.596
t14	0	240.29	0	212.452	13.438	200.996	11.881	177.711
t15	0	193.12	0	170.746	12.597	156.287	11.137	138.181
t16	0	160.61	0	142.003	14.551	118.063	12.865	104.385
t17	0	146.318	0	129.367	15.522	100.931	13.724	86.142
t18	0	127.224	0	112.485	18.18	74.065	16.074	65.485
t19	75.188	70.744	76.582	72.055	75.188	70.744	76.582	72.055
t20	72.389	65.986	73.731	67.209	72.389	65.986	73.731	67.209
t21	65.517	64.566	66.731	65.763	65.517	64.566	66.731	65.763
t22	49.215	64.043	50.128	65.23	49.215	64.043	50.128	65.23
t23	20.127	66.652	20.501	67.887	20.127	66.652	20.501	67.887
t24	71.365	11.139	72.688	11.346	71.365	11.139	72.688	11.346
	Operational cost No DR		Operational cost with DR		Operational cost No DR		Operational cost with DR	
	72379.940 $		70230.950 $		81471.210 $		78307.862 $	

References

[1] M. Mohammadi, Y. Noorollahi, B. Mohammadi-ivatloo, H. Yousefi, "Energy hub: From a model to a concept – A review," Renewable and Sustainable Energy Reviews, 80: 1512–1527, 2017.

[2] A. Yazdaninejadi, A. Hamidi, S. Golshannavaz, F. Aminifar, S. Teimourzadeh, "Impact of inverter-based DERs integration on protection, control, operation, and planning of electrical distribution grids," Electricity Journal, 32(6): 43–56, 2019.

[3] A. Yazdaninejadi, M. Farsadi, T. Sattarpour, "Optimal placement and operation of BESS in a distribution network considering the net present value of energy losses cost," ELECO 2015 - 9th International Conference on Electrical and Electronics Engineering , 2016.

[4] M. Nikzad, A. Samimi, "Responsive load model integration with SCUC to design time-of-use program," Journal of Electrical and Computer Engineering Innovations, 6(2): 217–226, 2019.

[5] M. Majidi, S. Nojavan, K. Zare, "A cost-emission framework for hub energy system under demand response program," Energy, 134: 157–166, 2017.

[6] T. Krause, G. Andersson, K. Fröhlich, A. Vaccaro, "Multiple-energy carriers: Modeling of production, delivery, and consumption," Proceedings of the IEEE, 99(1): 15–27, 2011.

[7] Y. Wang, N. Zhang, Z. Zhuo, C. Kang, D. Kirschen, "Mixed-integer linear programming-based optimal configuration planning for energy hub: Starting from scratch," Applied Energy, 210: 1141–1150, 2018.

[8] M. Geidl, G. Andersson, "Optimal power flow of multiple energy carriers," IEEE Transactions on Power Systems, 22(1): 145–155, 2007.

[9] A. Santhosh, A. M. Farid, K. Youcef-Toumi, "Real-time economic dispatch for the supply side of the energy-water nexus," Applied Energy, 122: 42–52, 2014.

[10] T. Ma, J. Wu, L. Hao, "Energy flow modeling and optimal operation analysis of the micro energy grid based on energy hub," Energy Conversion and Management, 133: 292–306, 2017.

[11] T. Ma, J. Wu, L. Hao, D. Li, "Energy flow matrix modeling and optimal operation analysis of multi energy systems based on graph theory," Applied Thermal Engineering, 146: 648–663, 2019.

[12] S. Derafshi Beigvand, H. Abdi, M. La Scala, "Optimal operation of multicarrier energy systems using Time Varying Acceleration Coefficient Gravitational Search Algorithm," Energy, 114: 253–265, 2016.

[13] A. Pepiciello, A. Vaccaro, M. Mañana, "Robust optimization of energy hubs operation based on extended affine arithmetic," Energies, 12(12): 2420, 2019.

[14] R. Z. Ríos-Mercado, C. Borraz-Sánchez, "Optimization problems in natural gas transportation systems: A state-of-the-art review," Applied Energy, 147. Elsevier Ltd: 536–555, 2015.

[15] D. De Wolf, Y. Smeers, "The Gas Transmission Problem Solved by an Extension of the Simplex Algorithm," Management Science, 46(11): 1454–1465, 2000.

[16] J. Wang, Y. Sun, Z. Xu, J. Xiong, "Optimization Dispatch of Integrated Natural Gas and Electricity Energy System under the Mode of Electricity-Orientated," in Proc. iSPEC 2019 - 2019 IEEE Sustainable Power and Energy Conference: Grid Modernization for Energy Revolution, Proceedings: 584–589, 2019.

[17] A. Martinez-Mares, C. R. Fuerte-Esquivel, "A unified gas and power flow analysis in natural gas and electricity coupled networks," IEEE Transactions on Power Systems, 27(4): 2156–2166, 2012.

[18] S. D. Beigvand, H. Abdi, M. La Scala, "A general model for energy hub economic dispatch," Applied Energy, 190: 1090–1111, 2017.

[19] M. Batić, N. Tomašević, G. Beccuti, T. Demiray, S. Vraneš, "Combined energy hub optimisation and demand side management for buildings," Energy and Buildings, 127: 229–241, 2016.

[20] A. Dolatabadi, B. Mohammadi-Ivatloo, M. Abapour, S. Tohidi, "Optimal Stochastic Design of Wind Integrated Energy Hub," IEEE Transactions on Industrial Informatics, 13(5): 2379–2388, 2017.

[21] M. HojatyDana, M. AlizadehPahlavani, "Control-Strategies-for-Performance-Assessment-of-an- Autonomous Wind Energy Conversion System," Journal of Electrical and Computer Engineering Innovations, 2(1): 15–20, 2014.

[22] M. Schulze, L. Friedrich, M. Gautschi, "Modeling and optimization of renewables: Applying the energy hub approach," 2008 IEEE International Conference on Sustainable Energy Technologies, ICSET 2008: 83–88, 2008.

[23] M. Geidl, P. Favre-Perrod, B. Klöckl, G. Koeppel, "A greenfield approach for future power systems," 41st International Conference on Large High Voltage Electric Systems 2006, CIGRE 2006, 2006.

[24] P. Favre-Perrod, M. Geidl, B. Klöckl, G. Koeppel, "A vision of future energy networks," in Proceedings of the Inaugural IEEE PES 2005 Conference and Exposition in Africa, 2005: 13–17, 2005.

[25] A. Soroudi, T. Amraee, "Decision making under uncertainty in energy systems: State of the art," Renewable and Sustainable Energy Reviews, 28: 376–384, 2013.

[26] M. Nikzad, A. Samimi, "Integration of Optimal Time-of-Use Pricing in Stochastic Programming for Energy and Reserve Management in Smart Micro-grids," Springer,2020.

[27] R. Hemmati, H. Saboori, and P. Siano, "Coordinated short-term scheduling and long-term expansion planning in microgrids incorporating renewable energy resources and energy storage systems," Energy, vol. 134, pp. 699–708, 2017.

[28] J. M. Nahman, D. M. Perić, "Radial distribution network planning under uncertainty by applying different reliability cost models," International Journal of Electrical Power and Energy Systems, 117: 105655, 2020.

[29] S. J. Ben Christopher, M. Carolin Mabel, "A bio-inspired approach for probabilistic energy management of micro-grid incorporating uncertainty in statistical cost estimation," Energy, 203: 117810, 2020.

[30] S. Xie, Z. Hu, J. Wang, "Scenario-based comprehensive expansion planning model for a coupled transportation and active distribution system," Applied Energy, 255: 113782, 2019.

[31] N. Bazmohammadi, A. Anvari-Moghaddam, A. Tahsiri, A. Madary, J. C. Vasquez, J. M. Guerrero, "Stochastic Predictive Energy Management of Multi-Microgrid Systems," Applied Sciences, 10(14): 4833, 2020.

[32] Y. Zhang, F. Meng, R. Wang, B. Kazemtabrizi, J. Shi, "Uncertainty-resistant stochastic MPC approach for optimal operation of CHP microgrid," Energy, 179: 1265–1278, 2019.

[33] W. Liang, K. C. Li, J. Long, X. Kui, A. Y. Zomaya, "An Industrial Network Intrusion Detection Algorithm Based on Multifeature Data Clustering Optimization Model," IEEE Transactions on Industrial Informatics, 16(3): 2063–2071, 2020.

[34] X. Li, Y. Li, L. Liu, W. Wang, Y. Li, Y. Cao, "Latin Hypercube Sampling Method for Location Selection of Multi-Infeed HVDC System Terminal," Energies, 13(7): 1646, 2020.

[35] A. Tabandeh, A. Abdollahi, M. Rashidinejad, "Transmission Congestion Management Considering Uncertainty of Demand Response Resources' Participation," Journal of Electrical and Computer Engineering Innovations, 3(2): 77–88, 2015.

[36] U. Mukherjee, S. Walker, A. Maroufmashat, M. Fowler, A. Elkamel, "Development of a pricing mechanism for valuing ancillary, transportation and environmental services offered by a power to gas energy system," Energy, 28: 447–462, 2017.

[37] A. Najafi-Ghalelou, S. Nojavan, K. Zare, B. Mohammadi-Ivatloo, "Robust scheduling of thermal, cooling and electrical hub energy system under market price uncertainty," Applied Thermal Engineering, 149: 862–880, 2019.

[38] U. Güvenç, B. Özkaya, H. Bakir, S. Duman, O. Bingöl, "Energy Hub Economic Dispatch by Symbiotic Organisms Search Algorithm," in Lecture Notes on Data Engineering and Communications Technologies, 43: 375–385, 2020.

[39] S. Galvani, S. Rezaeian Marjani, J. Morsali, M. Ahmadi Jirdehi, "A new approach for probabilistic harmonic load flow in distribution systems based on data clustering," Electric Power Systems Research, 176: 105977, 2019.

[40] A. Soroudi, "Power system optimization modeling in GAMS". Springer, Cham, 2017.

[41] Ibm, "IBM ILOG CPLEX Optimization Studio", 2012.

[42] Y. Zhu, "Power System Loads and Power System Stability". Springer, 2020.

[43] H. Aalami, G. R. Yousefi, M. Parsa Moghadam, "Demand response model considering EDRP and TOU programs," in Transmission and Distribution Exposition Conference: 2008 IEEE PES Powering Toward the Future, PIMS 2008, 2008.

[44] R. Aazami, K. Aflaki, M. R. Haghifam, "A demand response based solution for LMP management in power markets," International Journal of Electrical Power and Energy Systems, 33(5): 1125–1132, 2011.

[45] S. Das, A. Abraham, A. Konar, "Automatic clustering using an improved differential evolution algorithm," IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans, 38(1): 218–237, 2008.

[46] H. Hosseinnejad, S. Galvani, P. Alemi, "Optimal Probabilistic Scheduling of a Proposed EH Configuration Based on Metaheuristic Automatic Data Clustering," IETE Journal of Research: 1–23, 2020.

[47] Y. Zhu, D. Tian, F. Yan, "Effectiveness of Entropy Weight Method in Decision-Making," Mathematical Problems in Engineering, 2020.

[48] X. Li, K. Wang, L. Liuz, J. Xin, H. Yang, C. Gao, "Application of the entropy weight and TOPSIS method in safety evaluation of coal mines," in Procedia Engineering, 26: 2085–2091, 2011.

Biographies

Hadi Hosseinnejad received his B.S. degree from the University of Urmia, Urmia, Iran, in 2013, and his M.S. degree from the University of Islamic Azad University of Urmia, Urmia, Iran, in 2015. Currently, he is pursuing the Ph.D. degree in the School of Electrical Engineering, Islamic Azad University of Urmia, Urmia, Iran. His research interests are in optimal scheduling, energy hub management, and power systems include economic and reliability analysis.
Email: h.hosseinnejad@iaurmia.ac.ir

Sadjad Galvani received his B.S. degree from the University of Tabriz, Tabriz, Iran, in 2005, and his M.S. degree from the University of Zanjan, Zanjan, Iran, in 2007, and a Ph.D. degree in Electrical Engineering, from Urmia University, Urmia, Iran, in 2013. Currently, he is an assistant professor in the Department of Power Engineering, Faculty of Electrical and Computer Engineering, Urmia University, Urmia, Iran. His current research interests include probabilistic assessment of power systems, Facts included operation of power systems, reliable and secure operation of power systems.
Corresponding Author, Email: s.galvani@urmia.ac.ir

Payam Alemi was born in Tabriz, Iran, in 1982. He received his B.S. degree from the University of Tabriz, Tabriz, Iran, in 2005, and his M.S. degree from the Science and Research Branch, Islamic Azad University, Tehran, Iran, in 2008, and Ph.D. degree in Electrical Engineering, from Yeungnam University, Gyeongsan, Korea, in 2014. Then he joined Simon Fraser University, BC, Canada for his postdoctoral program until 2016. Currently he is an assistant professor in the department of electrical engineering, Islamic Azad University, Urmia Branch, Urmia, Iran. His current research interests include the control of multilevel power converters, power loss analysis for converters, LCL filters, machine drives and DC-DC converters.
Email: payamalemi@gmail.com

How to cite this paper:
H. Hosseinnejad, S. Galvani, P.Alemi, "An Efficient Configuration for Energy Hub to Peak Reduction Considering Demand Response Using Metaheuristic Automatic Data Clustering," Journal of Electrical and Computer Engineering Innovations, 8(2): 233-254, 2020

DOI: 10.22061/JECEI.2020.7240.375

URL: http://jecei.sru.ac.ir/article_1468.html

Journal of Electrical and Computer Engineering Innovations
(JECEI)
Journal homepage: http://www.jecei.sru.ac.ir

Research paper

Design of a Microstrip Dual-Band Bandpass Filter Using Novel Loaded Asymmetric Two Coupled Lines for WLAN Applications

R. Salmani, A. Bijari*, S.H. Zahiri

Deptartment of Electrical Engineering, Faculty of Electrical and Computer Engineering, University of Birjand, Birjand, Iran.

Article Info

Article History:
Received 23 October 2019
Reviewed 02 December 2019
Revised 08 January 2020
Accepted 09 April 2020

Keywords:
Dual-band filter
Bandpass filter
Microstrip lines
Coupled lines
Transmission zeros

*Corresponding Author's Email Address:
a.bijari@birjand.ac.ir

Abstract

Background and Objectives: Due to the rapid development in wireless communications, bandpass filters have become key components in modern communication systems. Among the microwave filter technologies, planar structures of microstrip line are chosen, due to low profile, weight, ease of fabrication, and manufacturing cost.
Methods: This paper designs and simulates a new microstrip dual-band bandpass filter. In the proposed structure, three coupled lines and a loaded asymmetric two coupled line are used. The design method is based on introducing and generating the transmission zeros in the frequency response of a wideband single-band filter. A wideband frequency response is obtained using the three coupled lines, and the transmission zeros are achieved using the novel loaded asymmetric two coupled lines.
Results: The proposed dual-band filter is designed and simulated on a Rogers RO3210 substrate for WLAN applications. Dimension of the proposed filter is 11.22 mm × 13.04 mm. The electromagnetic (EM) simulation is carried out by Momentum EM (ADS) software. Simulation results show that the proposed dual-band bandpass filter has two pass-bands at 2.4 GHz and 5.15 GHz with a loss of less than 1 dB for two pass-bands.
Conclusion: Among the advantages of this filter, low loss, small size, and high attenuation between the two pass-bands can be mentioned.

Introduction

Microwave integrated circuits (MICs) and radio frequency integrated circuits (RFICs) require a filter to eliminate interference, select band, attenuate harmonics, or eliminate the modulation distortion caused by active circuits used in communication transceivers. The microstrip transmission lines are one of the most common planar transmission lines due to their simple construction using conventional lithography processes and easy integration with ICs. Microwave filters are used as essential elements in each communication system for discriminating the frequency components of interest from the unwanted ones. These filters play an essential role in the desired performance of the transceiver systems.

Considering the advancements of wireless communication, the increase in bandwidth, and the development of new standards, small-sized microwave filters with excellent performance and low cost are required. Today, multiple band operation is considered to solve the insufficient capacity of the communication systems. Thus, new microwave filters should be able to operate in two or more non-harmonic frequency bands. With the development of wireless communication standards in the ultra-wideband context, filters that can operate in two or more frequency ranges, like IEEE806.16, IEEE806.11, CDMA, and GSM are required [1]. The components of the dual-band or multiple band bandpass filter can be used to meet this requirement. Almost all bands that are used for commercial purposes

are very close in terms of position and bandwidths. For example, WiFi, WiMAX, and GSM systems operate in the frequency bands of 0.9-1.8 GHz, 2.4-2.45 GHz, 3.5 GHz and 5.2-5.25 GHz. Since 1997 with the approval of using wireless local area networks (WLAN) for commercial purposes, applications of this technology have grown quickly. According to IEEE 802.11a/b/g, WLAN is applied in the frequency bands of 2.4-2.45 GHz, and 5.2-5.25 GHz. Considering the performance of dual-band and multiple-band filters in the stop-band, their size and construction cost, their design is very challenging. Since achieving excellent characteristics for close pass-bands is difficult, recently, various methods and structures have been proposed to develop the novel flat multiple-band filters, like designing the filter in classic form, using multi-mode resonators (MMR), and introducing and generating transmission zeros in the frequency response of a wideband single-band filter.

In [2], a dual-band bandpass filter using open-circuit and short-circuit loaded stubs has been designed and simulated. The proposed dual-band filter includes a second-order bandpass filter and a third-order bandpass filter, which are designed independently. The characteristics of the dual-band filter are obtained by combining two single-band bandpass filters that increases the filter dimension. Another limitation of the scheme is the short-circuit using via to ground. The bandwidth of the pass-bands can be controlled using impedance and length of the stubs. In [3], a novel microstrip dual-band filter with excessive loss in the stop-band has been introduced. The structure of the proposed filter includes coupled transmission lines and radial stubs. A dual-band bandpass filter using a novel microstrip dual-mode resonator based on a split-loop rectangular resonator with an open-circuit stub loaded has been presented in [4]. The proposed filter is based on the cross-coupling of a pair of modified resonators. The first higher order spurious mode is located at about 6 GHz, which limits the higher stopband width. In [5], a dual-band bandpass filter using a five-mode resonator has been proposed. The first three resonance modes are used for the first pass-band, and the two other modes are used for the second pass-band. Although the filter proposed has employed a five-mode resonator, using five open-circuit loaded stubs to achieve this goal increases the proposed filter's dimensions. A simple and effective method for designing the dual-band bandpass filters with high isolation and wide stop-band using open-circuit resonators loaded with one stub has been presented in [6]. The proposed structure is based on the conventional stub loaded resonator (SLR) that increases the filter dimension. In [7], two cells have been presented for implementation in passive circuits with a wide stop-band. Both cells include step impedance resonators and DGS structures. Based on these two cells, two dual-band bandpass filters have been designed and constructed. Using this cells and cascade them to achieve the characteristics of a dual-band filter increases the filter's dimension. Two single-band and dual-band microstrip bandpass filters with source and load loaded with dual-mode ring resonators based on two-layer structures have been presented in [8].

In this study, the proposed filter is introduced based on introducing and generating transmission zeros in the frequency response of the wideband single-band filter for WLAN applications. A wideband frequency response is realized using the three coupled lines, and the novel loaded asymmetric two coupled lines are used to generate the transmission zeros in the wideband frequency response.

Filter Design

A. Three coupled lines

If two or more transmission lines are very close to each other, the power is coupled between two lines due to interference of EM fields. Such lines are known as coupled transmission lines comprised of two conductors that are very close, although more conductors can be used. It is usually assumed that the coupled transmission lines are in TEM mode, valid for strip structures and microstrip structures.

Fig. 1 (a) shows a parallel symmetric three coupled microstrip line with distance s. If the TEM mode is considered, the electrical characteristics of the coupled lines can be described by measuring the effective capacitances between the lines and the propagation speed on the line.

According to the transmission theory, a transmission line can be modeled as a capacitance, inductance, and resistance. A three coupled structure supports three pseudo-TEM modes called a, b, and c [9]. To obtain the coupled structure's parameters, it is sufficient to obtain the capacitive matrix [C] per unit length.

The symmetric three coupled microstrip lines provide two transmission zeros at f=0 and f=2f_0. In this study, three coupled microstrip lines are used instead of the conventional two coupled lines, which increases the coupling and fractional bandwidth. In this case, the current of ports 1, 3, and 5 is zero, and its transmission matrix is described based on the width of the lines and their distance [9]. The equivalent circuit of the three coupled line is shown in Fig. 1 (b).

B. Introducing a New Resonator

In this study, a new resonator is presented using loaded asymmetric two coupled lines, as shown in Fig. 2.

This new open-circuit resonator comprises an asymmetric two coupled line with characteristic impedance of $Z_{0\pi}$ and Z_{0C} with an electrical length of ϑ_1

and a loaded line with characteristic impedance of Z_2, and electrical length of ϑ_2.

For the asymmetric two coupled structure, the relationship of the four-port network, its voltages and impedance parameters are represented in the following matrix form.

$$[V]_{4\times1} = [Z]_{4\times4}[I]_{4\times1} \quad (1)$$

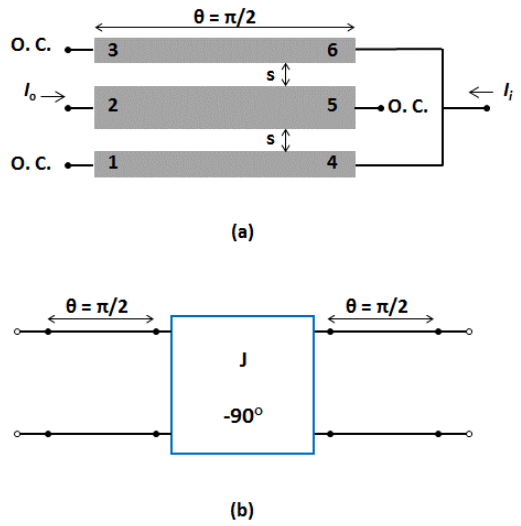


Fig. 1: (a) three coupled line, (b) its equivalent circuit.

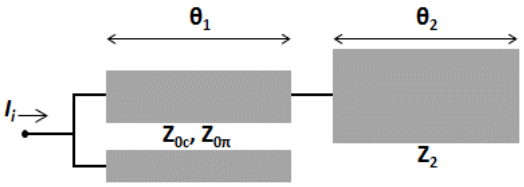


Fig. 2: The proposed new resonator.

The asymmetric two coupled line supports two pseudo-TEM propagation modes known as c and π. To obtain the parameters of the coupled structure, it is sufficient to obtain the capacitive matrix [C] per unit length.

The open-circuit impedance matrix of this four-port network can be obtained considering the superposition of modes c and π. The elements of the open-circuit impedance matrix [Z] are as follows [10]:

$$Z_{11} = Z_{33} = -j\left[\frac{Z_{01c}\cot\theta_c}{(1-R_c/R_\pi)} + \frac{Z_{01\pi}\cot\theta_\pi}{(1-R_\pi/R_c)}\right] \quad (1\text{-}2)$$

$$Z_{12} = Z_{21} = Z_{34} = Z_{43} = -j\left[\frac{Z_{01c}R_c\cot\theta_c}{(1-R_c/R_\pi)} + \frac{Z_{01\pi}R_\pi\cot\theta_\pi}{(1-R_\pi/R_c)}\right] \quad (2\text{-}2)$$

$$Z_{14} = Z_{41} = Z_{23} = Z_{32} = -j\left[\frac{Z_{01c}R_c}{(1-R_c/R_\pi)\sin\theta_c} + \frac{Z_{01\pi}R_\pi}{(1-R_\pi/R_c)\sin\theta_\pi}\right] \quad (3\text{-}2)$$

$$Z_{13} = Z_{31} = -j\left[\frac{Z_{01c}}{(1-R_c/R_\pi)\sin\theta_c} + \frac{Z_{01\pi}}{(1-R_\pi/R_c)\sin\theta_\pi}\right] \quad (4\text{-}2)$$

$$Z_{22} = Z_{44} = -j\left[\frac{Z_{01c}R_c^2\cot\theta_c}{(1-R_c/R_\pi)} + \frac{Z_{01\pi}R_\pi^2\cot\theta_\pi}{(1-R_\pi/R_c)}\right] \quad (5\text{-}2)$$

where ϑ_c and ϑ_π are the electric length of the asymmetric two coupled transmission line in mode c and mode π. The equations governing the electric length of the line are as follows: $\theta = \beta l$ that $\beta = \frac{2\pi}{\lambda_g}$ and $\lambda_g = \frac{c}{f_0\sqrt{\varepsilon_{re}}}$. Where l is the physical length of the line, β is the propagation constant, λ_g is the wavelength, and ε_{re} is the effective dielectric constant of the line. Z_{01C} and $Z_{01\pi}$ are the characteristic impedance of line 1, R_C and R_π are the parameters of the two excitation modes that are obtained using the capacitance per unit length. The Z_{01C}, $Z_{01\pi}$, R_C and R_π are defined in Appendix. Fig. 3 shows the asymmetric two coupled microstrip line used in the proposed resonator. This asymmetric coupled line operates as a band stop filter. The input is taken by connecting ports 1 and 2, and the output is taken from ports 3, while the ports 4 is open-circuit. "Equation (3)" describes the equivalent coupling line conditions:

$$I_4 = 0 \quad (1\text{-}3)$$

$$I_1 + I_2 = I_i \quad (2\text{-}3)$$

$$V_1 = V_2 = V_i \quad (3\text{-}3)$$

$$V_3 = V_o \quad (4\text{-}3)$$

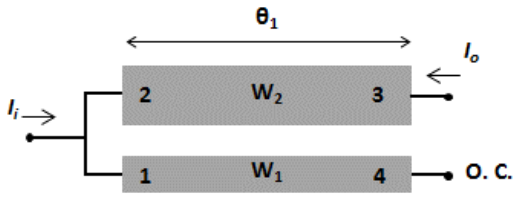


Fig. 3: The proposed structure using the asymmetric two coupled lines.

By applying the equivalent coupled line conditions in "(1)", we have:

$$V_i = Z_{11}I_1 + Z_{12}(I_i - I_1) + Z_{13}I_o \quad (1\text{-}4)$$

$$V_i = Z_{12}I_1 + Z_{22}(I_i - I_1) + Z_{14}I_o \quad (2\text{-}4)$$

$$V_o = Z_{13}I_1 + Z_{14}(I_i - I_1) + Z_{11}I_o \quad (3\text{-}4)$$

$$V_4 = Z_{14}I_1 + Z_{24}(I_i - I_1) + Z_{12}I_o \quad (4\text{-}4)$$

Therefore, by solving these equations, the impedance matrix of the equivalent two-port network is obtained using the following equations:

$$\begin{bmatrix} V_i \\ V_o \end{bmatrix} = \begin{bmatrix} Z'_{11} & Z'_{12} \\ Z'_{21} & Z'_{22} \end{bmatrix} \begin{bmatrix} I_i \\ I_o \end{bmatrix} \quad (5)$$

where

$$Z'_{11} = \frac{Z_{11}Z_{22} - Z_{12}^2}{Z_{11} - 2Z_{12} + Z_{22}} \quad (1\text{-}6)$$

$$Z'_{12} = Z'_{21} = \frac{Z_{11}Z_{14} - Z_{12}Z_{13} - Z_{12}Z_{14} + Z_{13}Z_{22}}{Z_{11} - 2Z_{12} + Z_{22}} \quad (2\text{-}6)$$

$$Z'_{22} = \frac{Z_{11}^2 - 2Z_{11}Z_{12} + Z_{11}Z_{22} - Z_{13}^2 + 2Z_{13}Z_{14} - Z_{14}^2}{Z_{11} - 2Z_{12} + Z_{22}} \quad (3\text{-}6)$$

The parameters of the above impedance matrix can be calculated using "(2)". By converting the impedance parameters of the dual network to transmission parameters of ABCD, we have [10]:

$$A = \frac{\cos\theta_C \cos\theta_\pi (R_C - R_\pi)}{R_C(1-R_\pi)\cos\theta_\pi - R_\pi(1-R_C)\cos\theta_C} \quad (1\text{-}7)$$

$$C = \frac{j}{R_C(1-R_\pi)\cos\theta_\pi - R_\pi(1-R_C)\cos\theta_C} \left[\frac{(1-R_C)^2\cos\theta_C}{R_C Z_{01\pi}\csc\theta_\pi} - \frac{(1-R_\pi)^2\cos\theta_\pi}{R_\pi Z_{01C}\csc\theta_C}\right] \quad (2\text{-}7)$$

$$\begin{aligned} D = {} & \{\cos\theta_C \cos\theta_\pi \left[R_\pi^2(1-R_C)^2 + R_C^2(1-R_\pi)^2\right] \\ & + R_C R_\pi \sin\theta_C \sin\theta_\pi \left[\frac{Z_{01C}}{Z_{01\pi}}(1-R_C)^2 + \frac{Z_{01\pi}}{Z_{01C}}(1-R_\pi)^2\right] \\ & - 2R_C R_\pi (1-R_C)(1-R_\pi)\} / (R_C - R_\pi) \\ & \left[R_C(1-R_\pi)\cos\theta_\pi - R_\pi(1-R_C)\cos\theta_C\right] \end{aligned} \quad (3\text{-}7)$$

$$B = \frac{1}{C}(AD - 1) \quad (4\text{-}7)$$

In "(7)," it is assumed that the electrical length of the line for both modes are equal ($\vartheta_1=\vartheta_C=\vartheta_\pi$). In Fig. 2, the transmission matrix of ABCD is obtained as the product of the two coupled line's transmission matrices and the open-circuit stub loaded. The transmission matrix of the equivalent dual line is represented with $[ABCD]_1$, which is extracted from "(7)". The transmission matrix of the loaded line is represented with $[ABCD]_2$. Thus:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix} \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix} \quad (8)$$

where

$$\begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix} = \begin{bmatrix} \cos\theta_2 & jZ_2\sin\theta_2 \\ \frac{j\sin\theta_2}{Z_2} & \cos\theta_2 \end{bmatrix} \quad (9)$$

Considering the definition of the transmission matrix as follows:

$$\begin{bmatrix} V_i \\ I_i \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} V_o \\ -I_o \end{bmatrix} \quad (10)$$

The input impedance of the equivalent network of Fig. 2, is calculated using the following equation, assuming that the two lines are of the same length ($\vartheta_1=\vartheta_2\cong\pi/2$). It should be noted that the stub loaded is open-circuit (I_o=0).

$$\begin{aligned} Z_i = {} & -\{jR_\pi \cot\theta[Z_2\cos^2\theta(R_C - R_\pi) \\ & + R_C R_\pi \sin^2\theta(R_C Z_{01C} - R_\pi Z_{01\pi})] \\ & R_C Z_{01C} Z_{01\pi}(R_C - R_\pi)\} / \{[2R_C R_\pi Z_{01C}\csc\theta \\ & ((R_\pi^2 - R_\pi + \frac{1}{2})R_C^2 - R_C R_\pi^2 + \frac{1}{2}R_\pi^2)Z_{01\pi}\sin\theta \\ & + Z_2(R_C^2 R_\pi Z_{01C} + (-R_\pi^2 Z_{01\pi} + (-2Z_{01C} \\ & + 2Z_{01\pi})R_\pi - Z_{01\pi})R_C + R_\pi Z_{01C})(R_C - R_\pi)]\cos^2\theta \\ & + R_C^2 R_\pi^2((R_C^2 Z_{01C}^2 + R_\pi^2 Z_{01\pi}^2 - 2R_C Z_{01C}^2 - 2R_\pi Z_{01\pi}^2 \\ & + Z_{01C}^2 + Z_{01\pi}^2)\sin^2\theta - 2Z_{01C}Z_{01\pi}(R_\pi - 1)(R_C - 1))\} \end{aligned} \quad (11)$$

For resonance, when Z_{in} is zero, by calculating ϑ for assumed R_π and R_C, the transmission zeros' location can be obtained using the following equations:

$$\theta = \begin{cases} \tan^{-1}\left(\sqrt{\frac{2NR}{R+1}}\right) \\ \frac{\pi}{2} \end{cases} \quad (12)$$

where

$$R = \frac{Z_{01C}}{Z_{01\pi}} \quad (1\text{-}13)$$

$$N = \frac{Z_2}{Z_{01C}} \quad (2\text{-}13)$$

If is defined at the central frequency of f_0, then the relationship between R and location of the transmission zeros for different values of N is calculated using Fig. 4. In all cases, three transmission zeros resulting from length, quarter wavelength at the central frequency of f_0 are shown.

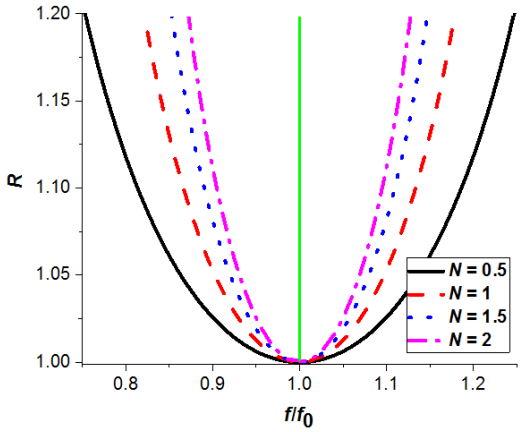


Fig. 4: Variations of the location of the transmission zeros vs. *R* for different values of *N*.

C. Implementation

In this paper, the initial design starts with a third-order elliptical bandpass filter and a ladder circuit equivalent to the lumped elements. Then, admittance inverters (*J*) are used to approximate the lumped elements to the distributed elements. An ideal admittance inverter is a two-port network in which the input admittance is equal to the inverse load admittance. Thus, it can be used to convert the series elements to parallel elements and vice versa. An admittance inverter can be constructed using a quarter-wave converter with proper characteristic impedance. The Schematic of the generalized proposed dual-band filter is shown in the following Fig. 5 (a). The proposed dual-band bandpass filter is shown in Fig. 5 (b). The proposed filter includes a symmetric three coupled microstrip line and loaded asymmetric two coupled microstrip line.

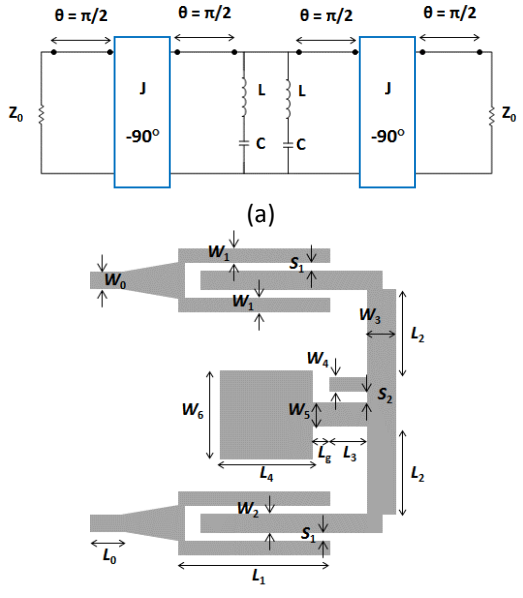


Fig. 5: (a) Schematic of the generalized proposed dual-band filter (b) The proposed dual-band filter.

Results and Discussion

The proposed dual-band microstrip filter shown in Fig. 5 (b), is designed and simulated on a Rogers RO3210 substrate with the thickness of 0.64 mm. The dielectric constant is 10.2, and the loss tangent is 0.0027. The proposed structure is simulated and optimized using momentum in ADS, as shown in Fig. 6. The dimensions after optimization are list in Table 1.

Table 1: Dimentsions of the proposed filter (milimeter)

W_0	W_1	W_2	W_3	W_4
0.58	0.33	0.27	1.05	0.20
W_5	W_6	S_1	S_2	L_0
1.80	4.85	0.12	0.10	0.75
L_1	L_2	L_3	L_4	L_g
7.65	4.75	0.32	4.80	0.50

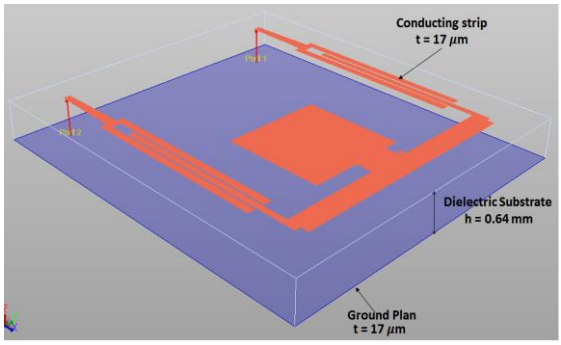


Fig. 6: Simulation of the proposed filter in Momentum ADS.

The frequency response of the proposed filter is shown in Fig. 7.

The proposed filter includes two pass-bands at the central frequencies of 2.4 GHz and 5.15 GHz with fractional bandwidths of 22.9% and 15.5%. The maximum insertion losses and the return losses in the first band are 0.5 dB and 11 dB, and they are 1 dB and 9 dB in the second band.

Also, there is one transmission zero between the two pass-bands at the frequency of 3.68 GHz with maximum attenuation of 49.5 dB.

Dimension of the proposed filter is 11.22 mm × 13.04 mm or 0.23 λ_g × 0.27 λ_g, where λ_g is the wavelength of the 50-Ω microstrip line over the substrate at central frequency of the first passband (2.4 GHz).

The input and output of the proposed filter are terminated with an impedance of Z_0=50 Ω. Therefore, the width of the feedlines (W_0=0.58 mm) are designed to provide matching with the characteristic impedance of 50 Ω.

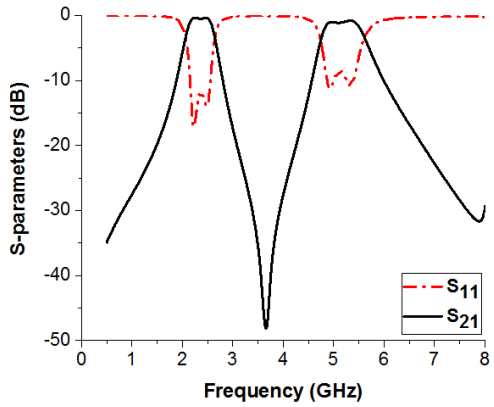


Fig. 7: Frequency response of the proposed filter.

In the following figure, the frequency response of the proposed filter is evaluated by the termination resistors of 25 Ω and 75 Ω. As can be seen, any mismatch at the circuits' input and output connected to the feed lines of the proposed filter can degrade the insertion loss at two frequency bands, particularly the second passband. As can be seen in Table 2, the proposed filter exhibits better insertion loss at the central frequencies, particularly at the first passband. Although the insertion loss of the filter in [2] is almost the same as the proposed filter, but its dimensions is 24.4 mm×17.5 mm, which is larger than the proposed filter.

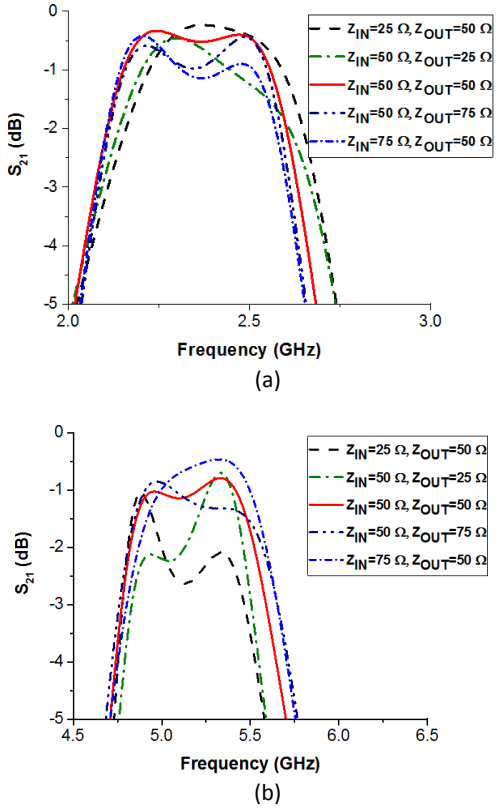


Fig. 8: Effect of termination resistors mismatch on the frequency response of the proposed filter at the (a) lower band (b) upper band.

Conclusion

This study presents the design, simulation, and optimization of a flat microstrip dual-band filter with symmetric three coupled lines and asymmetric two coupled lines. The proposed filter is designed for wireless local area networks (WLAN) with two pass-bands at the frequencies of 2.4 GHz and 5.15 GHz. The proposed filter has a maximum insertion loss of 0.5 dB and 1 dB and a return loss of 11 dB and 9 dB for the first and second pass-bands.

The proposed filter is designed and modeled systematically. This filter can also be implemented on a Rogers substrate with small dimensions of 11.22 mm × 13.04 mm or 0.23 λg × 0.27 λg, λg is the wavelength at the central frequency of the first pass-band.

The proposed filter could thus be a good choice for multiband receivers.

Author Contributions

R. Salmani and A. Bijari developed the theoretical idea and performed the analytic calculations. R. Salmani carried out the simulations. All authors discussed the results and contributed to the final manuscript. A. Bijari and S. H. Zahiri supervised the project.

Acknowledgment

We thank our colleagues from the university of Birjand who provided insight and expertise that greatly assisted the research. We thank M. Forouzanfar assistance for comments that greatly improved the manuscript.

Conflict of Interest

The author declares that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancy have been completely observed by the authors.

Abbreviations

WLAN	Wireless Local Area Network
MICs	Microwave Integrated Circuits
RFICs	Radio Frequency Integrated Circuits
CDMA	Code Division Multiple Access
GSM	Global System for Mobile
WiFi	Wireless Local Area Network Product Based on the IEEE 802.11
WiMAX	Worldwide Interoperability for Microwave Access
MMR	Multi Mode Resonator
EM	Electromagnetic
TEM	Transverse Electromagnetic mode

Table 2: Comparison of the proposed filter with its dual-band counterparts

Ref.	f_{01} (GHz)	f_{02} (GHz)	IL_{01} (dB)	IL_{02} (dB)	RL_{01} (dB)	RL_{02} (dB)	Size ($\lambda_g \times \lambda_g$)	Size(mm×mm)
[2]	2.4	5.2	0.3	0.7	22.1	20.8	0.28 × 0.20	24.4×17.5
[4]	2.12	3.91	0.92	2.11	17.3	15.4	0.24 × 0.18	26.4×20.3
[6]	3.5	5.25	1.87	2.33	>20	>20	0.459 × 0.323	27×19
[7] Filter A	3.17	3.91	1.76	1.63	>20	>20	0.389 × 0.177	27×12.5
[7] Filter B	3.16	3.90	1.87	1.67	>18	>20	0.385 × 0.213	26.9×14.75
[11] Filter A	1.8	5.8	1.33	1.7	21	13	0.23×0.17	28×20.5
[11] Filter B	2.4	5.8	1.35	1.97	17	15	0.39×0.25	35×22.5
[12]	2.82	3.21	1.9	1.7	21.6	16.1	2.76 × 1.30	21.4×10.1
[13]	3.78	4.82	1.38	1.82	14	33	0.16 × 0.31	-
This work	2.4	5.15	0.5	1	11	9	0.23 × 0.27	11.22×13.04

Appendix

The asymmetric two coupled line supports two pseudo-TEM propagation modes known as c and π. To obtain the parameters of the coupled structure, it is sufficient to obtain the capacitive matrix [C] per unit length. The phase velocity $v_{C,\pi}$ and the voltage ratio of the two lines $R_{C,\pi}$ are obtained from the following relations:

$$v_{c,\pi} = [\frac{D_1 + D_2}{2} \pm \frac{1}{2}\sqrt{(D_1 - D_2)^2 + 4E_1E_2}\,]^{-1/2} \quad (1\text{-}14)$$

$$R_{c,\pi} = \frac{1}{2E_1}[(D_2 - D_1) \pm \sqrt{(D_2 - D_1)^2 + 4E_1E_2}\,] \quad (2\text{-}14)$$

where

$$D_1 = (C_{11}C_{022} - C_{12}C_{012})/(v_0^2 \det[C_0]) \quad (1\text{-}15)$$

$$D_2 = (C_{22}C_{011} - C_{12}C_{012})/(v_0^2 \det[C_0]) \quad (2\text{-}14)$$

$$E_1 = (C_{12}C_{022} - C_{22}C_{012})/(v_0^2 \det[C_0]) \quad (3\text{-}14)$$

$$E_2 = C_{12}C_{011} - C_{012}^2 \quad (4\text{-}14)$$

$$[C_0] = C_{011}C_{022} - C_{012}^2 \quad (6\text{-}14)$$

where v_0 is the phase velocity in free space. The impedance of the i-th line in the j-th mode Z_{0ij} can be obtained from the following equations:

$$Z_{01j} = \frac{1}{v_j(C_{11} + R_jC_{12})} \quad (1\text{-}16)$$

$$Z_{02j} = \frac{1}{v_j(C_{22} + R_j^{-1}C_{12})} \quad (2\text{-}16)$$

The elements of the open-circuit impedance matrix [**Z**] are given in "(2)".

References

[1] V.C. Benjin, Advances in Multi-Band Microstrip Filters, Cambridge university press, 2015.

[2] G. Liang, F. Chen, "A Compact Dual-Wideband Bandpass Filter Based on Open-/Short-Circuited Stubs," IEEE Access, 8: 20488-20492, 2020.

[3] H. Abasi, A.R. Hazeri, "Compact microstrip dual-narrowband bandpass filter with wider rejection band," Electromagnetics, 38(6): 390–401, 2018.

[4] H. Chen, K. Chen, X. Chen, "Planar dual-mode dual-band bandpass filter using a modified rectangular split-loop resonator loaded by an open-circuited stub," Journal of Electromagnetic Waves and Applications, 30: 1964-1973, 2016.

[5] C.F. Chen, G.Y. Wang and J.J. Li, "Compact microstrip dual-band bandpass filter and quad-channel diplexer based on quint-mode stub-loaded resonators," IET Microwaves, Antennas & Propagation, 12: 1913-1919, 2018.

[6] Y. Xie, F.C. Chen, Z. Li, "Design of Dual-Band Bandpass Filter with High Isolation and Wide Stopband," IEEE Access, 5: 25602-25608, 2017.

[7] Y. Rao, H.J. Qian, B. Yang, R.G. Garcia, X. Luo, "Dual-Band Bandpass Filter and Filtering Power Divider with Ultra-Wide Upper Stopband Using Hybrid Microstrip/DGS Dual-Resonance Cells," IEEE Access, 8: 23624-23637, 2020.

[8] Q. Liu, D. Zhang, J. Zhang, D. Zhou, N. An, "Compact single- and dual-band bandpass filters with controllable transmission zeros using dual-layer dual-mode loop resonators," 14: 522-531, 2020.

[9] R. Schwindt, C. Nguyen, "Spectral Domain Analysis of Three Symmetric Coupled Lines and Application to a New Bandpass Filter," IEEE Transactions on Microwave Theory and Techniques, 42(7): 1183-1189, 1994.

[10] C. Nguyen, K. Chang, "On the Analysis and Design of Spurline Bandstop Filters," IEEE Transactions on Microwave Theory and Techniques, 33(12): 1416-1421, 1985.

[11] Z. C. Zhang, Q. X. Chu, F. C. Chen, "Compact Dual-Band Bandpass Filters Using Open-/Short-Circuited Stub-Loaded λ/4 Resonators," IEEE Microwave and Wireless Components Letters, 25(10): 657-659, 2015.

[12] R. Gomez-Garcia, J.-M. Munoz-Ferreras, W. Feng, D. Psychogiou, "Balanced symmetrical quasi-reflectionless single-and dual-band band- pass planar filters," IEEE Microwave Wireless Component Letter, 28(9): 798-800, 2018.

[13] L. T. Wang, Y. Xiong, L. Gong, M. Zhang, H. Li, X.-J. Zhao, "Design of dual-band bandpass filter with multiple transmission zeros using transversal signal interaction concepts," IEEE Microwave Wireless Component Letter, 29(1): 32-34, 2019.

Biographies

Reza Salmani received the BS degrees, in Telecommunications engineering and MSc degrees in Electronics engineering from University of Birjand, Iran, in 2014, and 2016, respectively. Currently, he is PhD Student, University of Birjand, Iran. Research interests: Microwave filters.

Abolfazl Bijari was born in Birjand, Iran in 1982. He received M.S. and Ph.D. in Electronics Engineering from Ferdowsi University of Mashhad (FUM), Iran in 2007 and 2013, respectively. He also took part in a year joint collaboration at the Synchrotron Light Research Institution (SLRI), in 2011 where he worked on LIGA-based micromechanical resonators. His research interest includes RF-MEMS, and RF circuit design for wireless communications. He is currently an Assistant Professor of Electrical Engineering at the University of Birjand.

Seyed Hamid Zahiri received the BS, MSc and PhD degrees in Electronics Engineering from Sharif University of Technology, Tehran, Iran, Tarbiat Modarres University, Tehran, and Ferdowsi University of Mashhad, Iran, in 1993, 1995, and 2005, respectively. Currently, he is a Professor in the Department of Electronics Engineering, University of Birjand, Iran. His research interests include pattern recognition, swarm intelligence algorithms, AI, and evolutionary algorithms.

How to cite this paper:
R. Salmani, A. Bijari, S.H. Zahiri, "Design of a microstrip dual-band bandpass filter using novel loaded asymmetric two coupled lines for WLAN applications," Journal of Electrical and Computer Engineering Innovations, 8(2): 255-262, 2020.

DOI: 10.22061/JECEI.2020.7250.376

URL: http://jecei.sru.ac.ir/article_1469.html

Journal of Electrical and Computer Engineering Innovations
(JECEI)
Journal homepage: http://www.jecei.sru.ac.ir

Research paper

Using Machine Learning Methods for Automatic Bug Assignment to Developers

M. Yousefi[1], R. Akbari[2,*], S. M. R. Moosavi[3]

[1] *E-Learning College, Shiraz University, Shiraz, Iran.*

[2]*Department of Computer Engineering and Information Technology, Shiraz University of Technology, Shiraz, Iran.*

[3]*Department of Computer Science, Engineering, and IT, Shiraz University, Shiraz, Iran.*

Article Info

Article History:
Received 07 September 2019
Reviewed 04 November 2019
Revised 12 December 2019
Accepted 12 March 2020

Keywords:
Automatic bug assignment
Bug reports
Bug clustering
Similarity criteria

[*]Corresponding Author's Email Address:
akbari@sutech.ac.ir

Abstract

Background and Objectives: It is generally accepted that the highest cost in software development is associated with the software maintenance phase. In corrective maintenance, the main task is correcting the bugs found by the users. These bugs are submitted by the users to a Bug Tracking System (BTS). The bugs are evaluated by the bug triager and assigned to the developers to correct them. To find a related developer to correct the bug, recent developers' activities and previous bug fixes must be examined. This paper presents an automated method to assign bugs to developers by identifying similarity between new bugs and previously reported bug reports.
Methods: For automatic bug assignment, four clustering techniques (i.e. Expectation-Maximization (EM), Farthest First, Hierarchical Clustering, and Simple Kmeans) are used where a tag is created for each cluster that indicates an associated developer for bug correction. To evaluate the quality of the proposed methods, the clusters generated by the methods are compared with the labels suggested by an expert triager.
Results: To evaluate the performance of the proposed method, we use real-world data of a large scale web-based system which is stored in the BTS of a software company. To select the appropriate algorithm for the clustering, the outputs of each clustering algorithm are compared to the labels suggested by the expert triager. The algorithm with closer output to the expert opinion is selected as the best algorithm. The results showed that EM and FarthestFirst clustering algorithms with 3% similarity error have the most similarity with the expert opinion.
Conclusion: the results obtained by the algorithms show that we can successfully apply them for bug assignment in real-world software development environments.

Introduction

Software systems enter into the maintenance phase after delivery to the customer and evolve over time. Software is constantly changing due to new changes needed by the customer and fixing possible bugs. Much of the cost of software development is spent on maintenance. Since software bugs are inevitable, it is imperative to assign the bug to a proper developer.

When a bug is reported in the software, the bug must be triaged. Bug triage is an important process in the software maintenance phase and has a major impact on software quality [1]. In the triage process, the person known as the triager, examines the accuracy of the reported bug. Valid bugs are then assigned to a developer to be fixed. The traditional and manual triage

process is time-consuming and costly and imposing more cost on the project [2].

In large-scale software projects, due to a large number of developers and the possibility that they may work parallel in various project modules, finding the appropriate developer is a difficult task and it is time-consuming and inaccurate to make the necessary checks [1], [2]. For example, the large number of bug reports or the wrong assignment of a bug slows down the debugging process. In this case, automatic bug assignment and clustering of bugs based on their similarities can make the triage and bug assignment more accurate and faster. Bugs in large software systems are maintained in BTS [3].

In large software projects, 50 to 60 bug reports are saved daily in the BTS [4]. As an example, for the Eclipse project, an average of 37 bug reports is logged daily in the BTS, which requires 3 person-hours per day for manual bug triage [5].

According to the study reported by Jeong et al., 44% of bugs have been assigned to the wrong developer after the first assignment [1]. To cope with this problem, in recent years, different types of methods have been proposed by authors [2]-[6]. These researchers were aimed to automate the bug triaging process. Some of the bug triage approaches are based on text categorization [2]. However, these methods suffer from poor quality reporting and cause to assign bugs to wrong developers [6], [7]. The main task in the bug assignment is to find the appropriate developer to fix the bug by analyzing the bug history that occurred in the software.

In this paper, an automated method for assigning the reported bug to the developer is presented in a closed source web-based software system. The method use clustering techniques to cluster the bugs. An expert opinion is used for accurate verification of the clustering algorithm and the outputs of each algorithm that are closest to the expert opinion are selected as the appropriate clustering algorithm. The main contributions of this paper are:

- Aggregating required data for bug triaging and assignment in a Closed Source Project (CSP).
- Using the proposed method for bug triaging in a real-world large scale web-based system.
- Adapting different machine-learning methods for data clustering and studying their performance for real-world data.

The remaining of this paper is organized as follows: the next section presents the previous works on bug triaging and bug assignment. The details of the proposed method is presented in Section “Methodology”. Section “Evaluation and Results” contains performance analysis and experimental results. Finally, conclusions and future works are given.

Related Work

In the maintenance phase, for bug triaging and bug assignment, many researchers use different information retrieval and machine learning methods to analyze textual sources in software repositories. More precisely, information retrieval and machine learning techniques have been extensively used by researchers to improve assigning bugs to developers. In this section, we review some of these methods for automatic bug assignment and bug triaging.

In [8], some bug assignment methods have been proposed. Different data sets and different input parameters have been used to evaluate the proposed method. According to this article, the number of different methods available for triage and bug correction confuses researchers. Therefore, in this paper, the work done to fix the bug is managed in a structured way. For this purpose, a structured combination of bug-solving methods is provided. Also, various aspects of bug correction are described and 6 related research questions in 5 dimensions are examined. To create infrastructure and organize bug assignment methods, 60 articles have been reviewed and classified. This study helps researchers to choose the right tools to fix the bug.

Limsettho et al. [9] presented a method for categorizing bug reports using topic modeling and two clustering algorithms. The proposed method has three phases. In the first phase, the bug reports are preprocessed and converted to topic vectors. These vectors are clustered in the second phase. Finally, each category of bugs is labeled.

Alenezi et al. proposed a method to reduce the bug triage time and automatic bug assignment to a related developer. They used Naive Bayes (NB) classifier to build a predictive model that can be used to assign a new bug report to a developer in the future. Five selection methods (LOR, X2, TFRF, MI, and DFS) have been used to reduce the size of the dimensions of terms and improve accuracy. This approach has two main steps. 1) A classification model is created using reduced terms to predict an experienced developer to fix newly reported problems. 2) Redistribute the load of overloaded developers. The evaluation was performed using four reported bugs from actual projects. Precision, recall, and F-score criteria were used to evaluate the performance of the classification [10]. The implementation of a recommendation system that was parallel and scalable and based on deep learning has been presented by Florea et al. [11]. Two deep learning categories have been used: Convolutional and Recurrent Neural Networks (CNN and RNN). The main theme of this article is not about running time, but about the scalability of the system on a cluster. This is measured using the speed criterion (the ratio of the sequential execution time to

the parallel execution time) and the parallel evaluation (speedup divided by the number of processors/cores) [11]. Shokripour proposed a method that uses textual information of the reported bugs in the bug repository to assign a new bug report to a developer. This method uses the term frequency-inverse document frequency (TF-IDF) term weighting technique. By using time metadata as an effective parameter in term weighting in term frequency-inverse document frequency, an attempt has been made to improve the automatic attribution of bug. The last time the term is used by the developer is used in the assignment [12].

In another work, Shokripour et al. presented a method based on Information Extraction (IE) techniques for bug assignment in large scale open source projects (OSP) [13]. The proposed method applied on three projects and more than 41% accuracies obtained.

Guo et al. presented a method based on convolution neural network (CNN) and developer activities for bug triaging [14]. They used CNN along with batch normalization and pooling to learn from the vectors generated by Word2vec. The performance of their method was evaluated on three open-source projects (OSPs). The bug assignment problem has been considered in [15] by employing programming keywords in the bug description as well as the recent expertise of developers. The authors applied their method on 93k bug-report assignments from 13 popular GitHub projects.

Zhang and Lee have proposed a method based on the combination of an experienced model and a probability model. First, a fixed bug that similar to new bug reports are extracted using the Smooth Unigram Model (SUM). Then an experienced model and a probability model based on similar bug reports are created. To create a probability model, social networking techniques are used to determine the relationship between developers from comments in the bug reports. The experience model is then created based on a series of project activity factors in the project such as the number of bugs which is fixed by the developer. Eventually, two models are combined and a developer rating is extracted that is used for new bug reports [16]. In other work, a machine learning-based approach was proposed that uses the nearest-neighbor algorithm to classify bug reports. The method consists of two components. The first component uses the VSM method with TF-IDF weighting to convert the fixed bug report text to the term vector space and determine the similarity of bug reports to new bug reports. The second component uses social network metrics to rank developers so that a ranking list is created based on the records of developer participation in discussing similar bug reports [17].

Kashiwa used mathematical programming for bug assignment [18]. He presented an optimization method called Release Aware and Prioritized Bug Fixing Task Assignment Optimization (RAPTOR). The purpose of this method is to mitigate the task concentration and increasing the number of bugs that developers can fix.

The application of ensemble methods has been studied by Goyal and Sardana in [19]. They used five ensemble methods called Bagging, Boosting, Majority Voting, Average Voting, and Stacking. For designing these ensembles, 25 different machine learning classifiers have been used by the authors. They applied these ensembles on three OSPs. Their results showed that the ensemble methods provide better performances in comparison with the base classifiers. In [20], an algorithm based on the Developer's Expertise Score (DES) for Bug Tossing Length (BTL) has been provided. The strategy is done in two steps: The first step is an offline process for obtaining a DES, which is calculated based on priority, adaptability, and average fixed time in developer activities. The online system process involves finding capable developers using three similarity calculation criteria (feature-based, cosine similarity, and Jacquard). The second step in the online process is to create points. Hit-ratio and reassignment accuracy are used to evaluate performance. In this method, the system is compared with ML-based debugging methods using three types of classification algorithms: Navies Bayes, Support Vector Machine (SVM), and C4.5 paradigms. By testing 41622 bug reports related to Mozilla, Eclipse, Netbeans, Firefox, and Freedesktop projects, the proposed method has an average accuracy of 89.49%, the precision is 89.53%, the recall rate is 89.42% and the F-score is 89.49%, which reduces BTL to 88.5%, which shows 20% improvement over existing technologies [20].

In [21], the main goal is to create a classifier to classify the reported bugs into two predefined classes: corrective report (defect fixing) and perfective report (major maintenance). This allows the maintainers to understand the bug more quickly when new bugs are reported and to provide the resources needed to fix the bug. For this purpose, the proposed method is based on a set of specific features that are based on the occurrence of specific keywords. This set is fed to some classification algorithms to create a classification model. The results of the proposed method are based on 3 different open source projects with an average accuracy of 93.1% with classification using the SVM classification algorithm [21].

The bug assignment problem in a CSP has been considered in [22], the goal was to reduce the bug assignment time to a developer with a related specialty that is reduced by tossing length. The development of such a technique is especially challenging for closed source projects. In this paper, a score is created to identify and rank an expert developer independent of the nature of the project. Two criteria are presented

based on developer expertise and bug importance score. These two criterion are calculated using information obtained from the components and content of the bug report. To validate the proposed method, the bugs that have been reported in a CSP developed by XYZ, pvt. Ltd has been used. The result obtained for the proposed method on the selected data set has been predicted with an accuracy of 88.9% .

In [23], a method for simultaneous bug triage was proposed for the developer and the development team using two-output neural network structure (called Dual DNN). This simultaneous is used using assignments made to the developer by team classes. A multi-label classification method has been used for two outputs for learning. A combination of exploratory labels that become a function of probability has been used. First, a two-step learning plan is used, in the first step of learning a part of the team is trained, and then the communication training between the team-developer and the developer-bug is done. The scheme is designed to encode team and developer relationships based on an organizational chart, which reinforces this model of organizational change because it can be adapted to role changes in an organization. A method called KSAP (K-nearest-neighbor search and heterogeneous proximity) was proposed by Zhang et al. to automatically assign a bug to the developer using historical bug reports and a heterogeneous network of bug repository [24]. When a new bug is reported, the bug is assigned to the developer in two phases. The first phase is to find similar bug reports to the new bug using the K-nearest-neighbor (KNN) method, and the second phase is to find developers who have participated in similar bugs using Heterogeneous proximity. An experiment on the Mozilla, Eclipse, Apache Ant, and ApacheTomcat 6 projects concluded that the KSAP method could improve the bug assignment recall between 7.5% and 32.25% compared to similar new methods [25].

Lee et al. reported that most previous studies focused only on OSPs and did not consider deep-learning techniques [25]. The Convolutional Neural Network (CNN) from the machine learning branch and Word2Vec from the word embedding branch has been used for automatic bug triage. The results obtained from the proposed method on the industrial project and open-source show the advantages of the approach. In fact, by using deep learning, the automatic assignment of a bug is performed on an industrial project. The performance advantages of the proposed method have been measured in comparison with human triage in terms of accuracy and simultaneous overhead. According to bug reports for industrial projects, we simulate the situations in which the proposed system is used and confirmed the effectiveness of the proposed system [25].

A two-phase method that used the Association Rule Mining (ARM) and X-Menas algorithm was proposed by Sharma and Singh for bug triaging [26]. In the first phase, the Apriori algorithm was used to predict the assignment of new bugs. The second phase used X-Means clustering along with ARM in each cluster. The performance of the proposed method was studied on some open source projects.

Mahendran proposed an approach that uses chart databases to calculate points for engineers and assign bugs to them [27]. This method is preferred over machine learning methods because there is no need to process of extracting, analyzing, or synchronizing data. The whole database for bug management can be in the graph database, and the method can be implemented directly on bug management tools. The proposed method controls the automatic assignment of errors along with workload balancing for engineers. Graph databases manage data internally as graphs and make relationships available as ready-made graphs in the database. It is possible to identify suitable maintenance engineers with queries without any specialized tools or extraction process.

Lee et al. proposed a two-phase method for cost-aware clustering of bug reports by employing the Genetic Algorithm (GA) [28] as an optimization algorithm. In the first phase of their method, a set of groups is created based on the similarities between bugs. The second phase constructs the clusters by grouping similar reports. The method was examined on the bug reports of Mozila's Firefox project.

A short survey of the previous methods is presented in Table 1. The second column shows the method used by the authors mentioned in column one. The third and fourth columns show the name and type of dataset used to evaluate the proposed methods.

As can be seen from Table 1, in most of the studies, data of OSPs have been used by researchers and there are a few works that have considered CSP data sets. The main OSPs that have been used in these studies are Eclipse, NetBeans, and Mozilla. Therefore, working on real-world data (or CSPs) and studying the applicability of the machine learning methods in this domain helps us to know if these methods are successful in CSPs or not. This fact encouraged us to study the bug-assignment problem in real-world environments. Also, previous works showed that in recent years different methods ranging from machine to deep learning, text mining and optimization have been used to cope with the bug assignment problem. It should be noted that in this article the emphasis is not on improving machine learning methods or other methods, but the main emphasis is on using these methods in the real world. Hence, some clustering algorithms have been used in their classical form.

Table 1: A survey on previous work, used methods and datasets

Ref.	Method	Dataset	Dataset type
Limsettho et al. (2016) [9]	Topic Modeling, EM, X-Means	HTTPClient, and JCR	OSP
Alenezi et al. (2013) [10]	Naive Bayes	Eclipse-SWT, Eclipse-UI, NetBeans, Maemo	OSP
Florea et al. (2017) [11]	CNN and RNN	Netbeans, Eclipse and Mozilla	OSP
Shokripour et al. (2015) [12]	ABA-Time-tf-idf	Eclipse, NetBeans, ArgoUML	OSP
Shokripour et al. (2012) [13]	IE methods	Eclipse, Mozilla, and Gnome	OSP
Guo et al. (2020) [14]	CNN	Eclipse, Mozilla and NetBeans	OSP
Sajedi-Badashian, and Stroulia (2020) [15]	Vocabulary and Time-aware Bug-Assignment (VTBA)	13 popular GitHub projects	OSP
Zhang, and Lee (2013) [16]	Unigram Model (UM)	Jboss, and Eclipse	OSP
Wu et al. (2011) [17]	KNN, expertise ranking	Mozilla Firefox	OSP
Kashiwa, Y. (2019) [18]	Mathematical programming	Mozilla Firefox, Eclipse, and GNU compiler collection (GCC)	OSP
Goyal, and Sardana (2019) [19]	Bagging, Boosting, Majority Voting, Average Voting, and Stacking	Mozilla Firefox, Open Office, and GNOME	OSP
Yadav et al. (2019) [20]	DES based online system	Mozilla, Eclipse, Netbeans, Firefox, and Freedesktop	OSP
Otoom et al. (2019) [21]	SVM	AspectJ, Tomcat, SWT	OSP
Yadav et al. (2018) [22]	A metric based method	XYZ, pvt. Ltd. India	CSP
Choquette-Choo et al. (2019) [23]	Dual DNN	Google Chromium project	OSP
Zhang et al. (2015) [24]	KSAP	Mozilla, Eclipse, Apache Ant, and ApacheTomcat 6	OSP
Lee, Sun-Ro, et al. (2017) [25]	CNN, Word Embedding	JDT, Platform, Firfox /A,B,C,D	OSP/CSP
Sharma and Singh (2016). [26]	ARM, X-Means	Thunderbird, Add-on SDK, and Bugzilla	OSP
Satish, and Mahendran (2018) [27]	Page ranking and graph databases	QT Framework	OSP
Lee et al. (2019) [28]	GA	Mozilla's Firefox	OSP

Methodology

This section presents the proposed method in detail. The proposed method is aimed to triage the bug report and assign it to an appropriate developer with acceptable speed when a bug is reported in a CSP. The proposed method uses clustering techniques to classify similar bug reports. To select the appropriate clustering algorithm, they are evaluated and the most appropriate algorithm is selected. Determining the similarity between bug reports is done by analyzing their context. Before discussing the proposed method more accurate, some of its advantages are as follows:

- In the real world, we usually face a lot of errors, and it is possible that many of the reported errors are of the same type and go into the fixed state without being checked and no assignment is made to them. The proposed method help the triager to mitigate this problem.
- The speed and accuracy of the bug assigned to the developer increases and developers who have to do the debugging are more accurately identified.
- We can manage all the bugs that affect a specific business or a particular software feature in a single cluster and get enough information from them.
- All bugs in the bug repository are categorized and grouped, making it easy to access and manage as well as reporting.

In software companies, the bug assignment is done by the bug triager manually. She/he checks the reported bugs and select the appropriate developer(s) to fix them. In the proposed method, we use the assignments proposed by the experts as our reference. Hence, the clustering algorithm that suggests the assignments that are closer to the assignments of the experts is preferred.

A. Description of the real-world case study

The proposed method is aimed to process the real data of a web-based system developed (we call it xyzSystem) in a software company. The data used here are the bug reports that have been stored in the BTS of the software company. The BTS maintains the bug reports of three software projects. These projects belong to a larger project. Three software work together to achieve a common business goal. One of the software mentioned as the main software receives online services from the other two software. The purpose of this web-based software is to manage corporate purchases.

B. Data Gathering

The users of the xyzSystem can report the bugs during working with the system. For this purpose, they log in to the ticketing system and send the bug report directly to the BTS. The BTS records are processed by the change control board and after validating the bug report, the appropriate developer is selected to fix the bug. This process is done manually. So, it is a time-consuming task. Automating this task helps the maintenance team to save time and cost and user satisfaction increases.

To apply the proposed method, the bug reports submitted by the users through a ticketing system is used. The bug reports are maintained in the BTS repository. For this purpose, the bug reports are extracted from the BTS using a wrapper, converted to the appropriate format, and arranged in an Excel file. The steps for preparing data can be seen in Fig. 1.

C. Overview of the Method

The steps of the proposed method can be seen in Fig. 1. Bug reports in software bug tracking systems are the information needed to start the proposed method. At the start of the process, the bug reports extracted from the bug tracking system repository are used as the input of the process. Next, the preprocessing step is applied

By applying the preprocessing steps to the bug reports, each bug report is converted to a term vector. After creating the term vector, the similarity between each vector is calculated. By creating the similarity matrix, we are ready to apply clustering algorithms. Finally, we apply tagging on the clusters.

D. Pre-Processing steps

As shown in Fig. 1, the description and summary of each bug report are extracted and used as the input of the preprocessing step. The steps of the pre-processing phase convert raw data into the useful data.

The content of a bug report (which is a bug summary and description) contains information such as time the bug occurred, the location of the bug, and the cause of the bug. At first, the tokenization operation is applied to

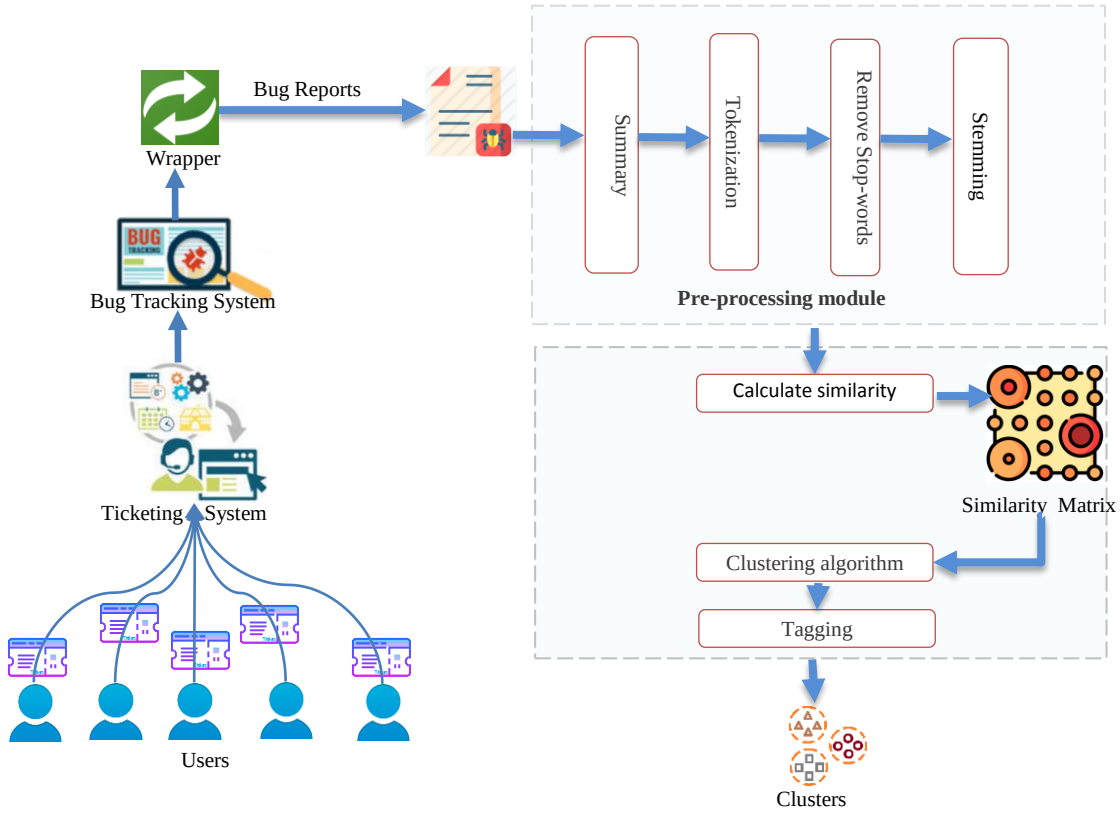


Fig. 1: Schematic diagram of the proposed method.

the extracted text of the bug report. In this way, the textual content of the bug report is converted to tokens. After that, Stop words are removed. The stemming (convert the word to base form) operation is performed on tokens. As an example of the operations in the preprocessing phase, Fig. 2 shows a sample bug report related to the xyzSystem, and the output of the pre-processing steps on the bug reports is shown in Fig. 3.

2019-04-08 10:13:36

Exception in org.xyzSystem.dominant.dao.core.nonPlanAllocation.INonPlanAllocationRepository.getAllGrid()

with cause = 'org.hibernate.exception.SQLGrammarException:

could not extract ResultSet' and exception = 'could not extract ResultSet;

Fig. 2: Summary of a bug report.

2019, 04, 08, 10, 13, 36, exception, org, xyzSystem, dominant, dao, core, nonplanallocation, inonplanallocationrepository, getallgrid, cause, org, hibernate, exception, sqlgrammarexception, could, extract, resultset, exception, could, extract, resultset

---

Fig. 3: Result of pre-processing steps.

E. Calculate Similarity

After applying the pre-processing step to the bug reports, the term vector of each report is calculated. The number of repetitions per term in each bug report is the term vector of that bug report. After calculating the term vector of bug reports, the similarity between the term vectors is calculated using Pearson's correlation coefficient. This is one of the most frequently used methods for calculating the data dependencies [29].

This coefficient is between 1 and -1 and is zero if no relationship exists between the two variables. The formula for calculating Pearson's correlation coefficient is as follows:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \quad (1)$$

In Pearson's formula, the values of x and y represent two vectors and the value of *n* represents the number of terms involved in calculating the term vector of bug reports. The similarity of each pairs of bug reports is calculated and the similarity matrix is generated between bug reporting vectors. As an example, Table 2 shows the similarity matrix of four typical bug reports in xyzSystem. According to Table 2, the values in the cells of the matrix indicate the similarity between bug reports in the related row and column.

Table 2: Similarity Matrix

	Bug1	Bug2	Bug3	Bug4
Bug1	1	-0.18786	0.099853	0.948872
Bug2	-0.18786	1	0.546667	-0.13032
Bug3	0.099853	0.546667	1	0.099853
Bug4	0.948872	-0.13032	0.099853	1

According to the permissible values of the correlation coefficient, if the similarity value obtained is near to 1 indicates the similarity and if is near to -1 indicates the non-similarity of the bug reports. According to the values in Table 2, reporting bug 1 and 4 with similarity values close to one are more similar, and error reporting 1 and 2 with similarity values near negative are non-similar. If we set the number of clusters to three by default, two bug reports 1 and 4 falls into one cluster and two other bug reports fall into separate clusters.

F. Clustering

The similarity matrix generated in the previous step is used as the input to the clustering algorithms. Clustering algorithms cluster the error reports in the matrix based on their similarity values. The clustering procedure in the proposed method detects related bug reports. Similar bug reports fall into a cluster. The clustering algorithms used here are: 1) EM, 2) Farthest First, 3) Hierarchical Clustering, and 4) Simple K-Means. The clustering algorithms implemented in Weka version 3.6.9 are used here.

G. Tagging

The purpose of clustering is to assign tags to objects that represent each object's membership in the cluster. These tags are keywords that indicate the identity of the content of the cluster.

After clustering, the bug reports should be tagged on the clusters. Usually, the suggestions for selecting a tag can be based on the number of repetitions of the term in the bug reports and the term that is most frequently repeated in the bug reports is selected as the cluster tag. In the proposed method, depending on the clustering issue associated with clustering bug reports and selecting the appropriate developer to fix the bug, each cluster is tagged with the developer specification that has fixed the bug.

Now, when a new bug report occurs in the xyzSystem, the steps of the proposed method are applied on it. First the pre-processing step is performed on the new bug report and finally based on the calculated similarity criteria for the new bug report against the clustered bugs, it is added to the most similar cluster. The new bug is assigned to the developer whose name is tagged on the target cluster. Also, the bug status is changed to “assigned”. The next section presents the test results in detail.

Evaluation and Results

It seems that the proposed method provides an efficient way for automating the bug triage process. In this section, the proposed method is tested with real dataset and the performance of the clustering algorithms is investigated.

A. Dataset

In order to evaluate the proposed method and to understand it more accurately and to verify the validity of the proposed method, experiments were performed on the real dataset of the CSP. The dataset used is the content of bugs that occurred within a given period in the xyzSystem and were fixed by developers with relevant knowledge. The content of the bugs is in a text format. The text file contains a summary of the error description with the exact date and time of the error, and the full address description of the class in which the error occurred, and the reason for the error in the summary of the error description. Fig. 2 shows an example of a brief description of a bug that contains the date and time the bug occurred, the location of the bug, and the cause of the bug. The bug text also contains complete bug descriptions that provide complete information about the bug occurring, and lists the classes inheriting from the original bug class, as well as the list of classes from which the bug class inherits. These items are used in the proposed method to extract the required features for clustering. Bugs that occur at different times on the system are stored in the software bug tracking system and from the time the bug was assigned to the developer until the bug is resolved, the history is stored in the system. All bugs resolved by a developer are considered as items in a cluster and that developer characteristic is tagged on the cluster. As an expert opinion, having a thorough knowledge of the system and the operating process of the system, five clusters were extracted from the bug tracking system. These five clusters containing 100, 50, 50, 50, 50 errors, respectively. So, we have 300 system bugs that were reviewed and resolved by five developers. The bugs are clustered by thoroughly examining the bug text, and each cluster is identified by a developer. That is, on each cluster, the name of the developer that should fix the bugs within that cluster is tagged.

B. Expert Opinion

In this work, we use the tags proposed by the expert triager for each bug report as our reference. According to the expert opinion, the tested dataset is extracted from the bug tracking system, with a total of 300 bugs reported during a specific period. A total of 100 bug fixes have been resolved by one developer, which is considered as a cluster, and the rest of the bugs have been evaluated in four 50-batch clusters by four other developers. Details of the bug reports of the xyzSystem suggested based on the opinions of the expert bug triager are shown in Table 3.

Table 3: Expert opinion specifications about the dataset and tag of each bug report

# of developers	5
# of clusters	5
# items in cluster	50, 100,50,50,50
# bug reports	300

C. Evaluation

After determining the optimal clusters based on the expert opinion, we are ready to evaluate the performance of the clustering methods. The accuracy of the existing clustering algorithms are measured and the algorithm with the near output to the expert opinion is determined. After that, the selected algorithm is used for clustering the new bug reports. Table 4 shows the output of four clustering algorithms. The second column shows the number of clusters. We set this number at 5, because we have 5 active developer in our case study. The third column shows the distribution of bugs in five clusters. For example, in EM algorithm we can see that the first and second cluster contains 49 and 101 bug reports respectively. Each of the three remaining clusters contains 50 bug reports. The fourth column shows the error rate of the corresponding algorithm. The error rate represent the number of bug reports that are incorrectly clustered.

Table 4: Evaluation of the clustering algorithms in terms of error rate and distribution of bug reports in clusters.

Algorithm	# of cluster	Cluster Instances	Error rate
EM	5	49 (16%) 101 (34%) 50 (17%) 50 (17%) 50 (17%)	0.3%
Farthest First	5	49 (16%) 101 (34%) 50 (17%) 50 (17%) 50 (17%)	0.3%
Hierarchical Cluster	5	100 (33%) 100 (33%) 50 (17%) 49 (16%) 1 (0%)	17%
Simple Kmeans	5	7 (2%) 0 (17%) 43 (14%) 101 (34%) 99 (33%)	19%

As can be seen from Table 4, the EM and FarthestFirst algorithms with 3% error rate are the most suitable algorithms for clustering. In EM and Farthest First algorithms, only one bug report is clustered incorrectly. It seems that one bug report from the first cluster is

determined as a member of the second cluster incorrectly. The Hierarchical Cluster and Simple Kmeans algorithms with 17%, and 19%, respectively obtain the next ranks. In hierarchical clustering, most of the bug reports that belong to the fifth cluster are incorrectly assigned to the first cluster. In simple Kmeans, we can see a different behavior where most of the members of the first cluster and all the members of the second cluster are recognized as the members of the fourth and fifth clusters. In general, all the algorithms examined in this work on real data have more that 80% accuracy. However, the EM and FarthestFirst algorithms have similar results, consistent with our expert opinion, and have competitive results. So we can use either of these two algorithms to cluster the bugs. The results of the two Hierarchical Cluster and Simple Kmeans placed at the third and fourth ranks respectively.

D. Applicability and limitations

The results showed that the proposed method based on the clustering techniques has the ability to generate good results for the xyzSystem. It seems that the proposed method is applicable to triage bugs in other CSPs. Usually, similar scenario is used by software companies to receive bug reports in the maintenance phase and triage the reported bugs. In this study, we have extracted 300 bug reports from the BTS to generate clustering models. However, several thousand bug reports available for large-scale OCPs such as Mozilla, Eclipse, etc. that have been used by authors in previous works. Hence, larger datasets in CSPs is recommended to be used in order to study the behavior of the clustering algorithm in such situations. Because the expert opinion is used for measuring the correctness of algorithms, there are factors that may have a negative impact on the algorithm. There must be assurance of the correctness of the expert opinion during different periods of time. Certainly, one of the limitations and challenges will be the inability to confirm the current evolution of expert opinion over time due to changes in the structure of projects and the updates of the development technologies. Another challenge is that past errors may not bear any resemblance to new errors. This is occurred due to possible changes in the project or the organization's focus on new projects and lack of investment in the support and development of past software systems. Another challenge is the change of the structure of the developer teams that can be a weakness in the assignment system because labels on clusters may be the names of developers who are blocked and no longer have a role in developing and supporting systems.

Conclusion

Identifying previously bug reports can reduce the cost of maintaining software. This paper proposed a method for clustering similar bug reports based on the similarity of the contextual content of the reported bugs. For each cluster, the corresponding developer's name is tagged. The calculation of similarity between bug reports is performed using Pearson's correlation coefficient. Four clustering algorithms have been evaluated by the considering the expert opinion. The appropriate algorithm with 3% error is selected for clustering. It seems that the proposed method and similar works can play an important role in maintenance phase to reduce the cost and speed up the bug fixing process. They can be used as an assistance for the bug triager or change control board in software development companies. However, more studies are needed to investigate different aspects of applying automation methods for bug triaging in CSPs.

Author Contributions

M. Yousefi collected the data and designed the experiments. A. Akbari carried out the data analysis. S. M. R. Moosavi and R. Akbari validated the results and wrote the manuscript.

Acknowledgment

The authors would like to thank Computer Engineering and IT Department of Shiraz University of Technology and Computer Science, Engineering, and IT Department of Shiraz University.

Conflict of Interest

The author declares that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancy have been completely observed by the authors.

Abbreviations

There is no abbreviations.

References

[1] G. Jeong et al., “Improving bug triage with bug tossing graphs,” in Proc. 7th Joint Meeting of the European Software Engineering Conference and the ACM SIGSOFT Symposium on the Foundations of Software Engineering: 111–120, 2009.

[2] J. Anvik, L. Hiew, G. Murphy, “Who should fix this bug?,” in Proc. 28th International Conference on Software Engineering: 361–370, 2006.

[3] D. Cubranic, C. Murphy, “Automatic bug triage using text categorization,” in Proc. Sixteenth International Conference on Software Engineering, Citeseer:92–97, 2004.

[4] H. Hu, H. Zhang, J. Xuan, W. Sun, “Effective bug triage based on historical bug-fix information,” in Proc. 25th International Symposium on Software Reliability Engineering: 122–132, 2014.

[5] J. Anvik, “Automating bug report assignment,” in Proc. 28th International Conference on Software Engineering, ACM: 937–940, 2006.

[6] J. Xuan, H. Jiang, Z. Ren, J. Yan, Z. Luo, “Automatic bug triage using semi-supervised text classification,” in Proc. Intl. Conf. Software Engineering & Knowledge Engineering: 209–214, 2010.

[7] N. Bettenburg, S. Just, A. Schro¨ter, C. Weiss, R. Premraj, T. Zimmermann, “What makes a good bug report?,” in Proc. 16th ACM SIGSOFT International Symposium on Foundations of software Engineering, ACM:308–318, 2008.

[8] A. Goyal, N. Sardana, "Analytical study on bug triaging practices," Jaypee Institute of Information Technology, Department of Computer Science and Engineering, Noida, UP, India, 2020.

[9] N. Limsettho, H. Hata, A. Monden, K. Matsumoto, "Unsupervised bug report categorization using clustering and labeling algorithm," International Journal of Software Engineering and Knowledge Engineering, 26(07): 1027-1053, 2016.

[10] M. Alenezi, M. Kenneth, S. Banitaan, "Efficient bug triaging using text mining journal of software," 8(9): 2185–2190, 2013.

[11] A.-C. Florea, J. Anvik, R. Andonie, "Parallel implementation of a bug report assignment recommender using deep learning," Conference Paper inLecture Notes in Computer Science, 2017.

[12] R. Shokripour, "A time-based approach to automatic bug report assignment," Journal of Systems and Software, 102: 109-122, 2015.

[13] R. Shokripour, Z.M. Kasirun, S. Zamani, J. Anvik, "Automatic bug assignment using information extraction methods," in Proc. International Conference on Advanced Computer Science Application and Technologies (ACSAT): 1-7, 2012.

[14] S. Guo et al., "Developer activity motivated bug triaging: via convolutional neural network," Neural Processing Letters, 51: 2589-2606, 2020.

[15] A. Sajedi-Badashian, E. Stroulia, "Vocabulary and time based bug-assignment: A recommender system for open-source projects," Software: Practice and Experience, 50(8): 1539- 1564, 2020.

[16] T. Zhang, B. Lee, "A hybrid bug triage algorithm for developer recommendation. Proceedings of the 28th Annual ACM Symposium on Applied Computing, SAC'13, ACM, NewYork, NY, USA: 1088–1094, 2013.

[17] W. Wu et al. "Drex:developer recommendation with k-nearest-neighbor search and expertise ranking," in Proc. the 2011 18th Asia-Pacific Software Engineering Conference, APSEC: 389, 2011.

[18] Y. Kashiwa, "RAPTOR: Release-aware and prioritized bug-fixing task assignment optimization," in Proc. 2019 IEEE International Conference on Software Maintenance and Evolution (ICSME): 629-633, 2019.

[19] A. Goyal, N. Sardana, "Empirical analysis of ensemble machine learning techniques for bug triaging," in Proc. 2019 Twelfth International Conference on Contemporary Computing (IC3): 1-6, 2019.

[20] A. Yadav , S. Singh, J. Su, "Ranking of Software developers based on expertise score for bug triaging," Information and Software Technology, 112: 1-17, 2019.

[21] A. Otoom et al."Automated classification of software bug reports," in Proc. the 9th International Conference on Information Communication and Management: 17–21, 2019.

[22] A. Yadav, D. Singh, "An information-theoretic approach for bug triaging," 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2018.

[23] C. Choquette-Choo et al. "A multi-label, dual-output deep neural network for automated bug triaging," 18th IEEE International Conference On Machine Learning and Applications (ICMLA), 2019.

[24] W. Zhang, S. Wang , Q. Wang, KSAP:An approach to bug report assignment using KNN search and heterogeneous proximity. Article in Information and Software Technology, 70: 68-84, 2015.

[25] S.-R. Lee, et al., "Applying deep learning based automatic bug triager to industrial projects," ESEC/FSE 2017: in Proc. the 2017 11th Joint Meeting on Foundations of Software Engineering, 926–931, 2017.

[26] M. Sharma, V.B. Singh, "Clustering-based association rule mining for bug assignee prediction," International Journal of Business Intelligence and Data Mining, 11(2): 130-150, 2016.

[27] S. C J, A. Mahendran, "Automated bug assignment in software maintenance using graph databases," International Journal of Intelligent Systems and Applications, 2: 27-36, 2018.

[28] J. Lee, D. Kim, W. Jung, "Cost-Aware clustering of bug reports by using a genetic algorithm," J. Inf. Sci. Eng., 35(1): 175-200, 2019.

[29] M.C. Abounaima et al., "The pearson correlation coefficient applied to compare multi-criteria methods: case the ranking problematic," in Proc. 2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET): 1-6, 2020.

Biographies

Mehran Yousefi received his BSc from Isfahan University of Technology. Also, he received his MSc from Shiraz University. His research interests are software engineering, program analysis, software security, reliability of software, AI systems, and Big data. He also has experience in developing software using python, java, php, and groovy.

Reza Akbari has a PhD in software engineering from Shiraz University. Currently, he is an associate professor at department of Computer Engineering and Information Technology of Shiraz University of Technology. His special fields of interest include software engineering in general, machine and deep learning, and optimization algorithms.

Mohammad Reza Moosavi received M.S. and Ph.D. in Software Engineering from Shiraz University, where he is currently an assistant professor. His research interests are Data Mining, Statistical Pattern Recognition and Distributed Systems. He also has teaching experiences especially in field of graph mining, formal methods and distributed systems.

How to cite this paper:
M. Yousefi, R. Akbari S.M.R. Moosavi, "Using machine learning methods for automatic bug assignment to developers," Journal of Electrical and Computer Engineering Innovations, 8(2): 263-272, 2020.

DOI: 10.22061/JECEI.2020.7212.370

URL: http://jecei.sru.ac.ir/article_1471.html

Journal of Electrical and Computer Engineering Innovations
(JECEI)
Journal homepage: http://www.jecei.sru.ac.ir

Research paper

A New Clustering Algorithm for Attributive Graphs through Information Diffusion Approaches

S. Kianian[1], S. Farzi[2,], H. Samak[2]*

[1]Faculty of Computer Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran.

[2]Faculty of Computer Engineering, K. N. Toosi University of Technology, Tehran, Iran.

Article Info

Article History:
Received 02 September 2019
Reviewed 07 November 2019
Revised 13 December 2019
Accepted 13 March 2020

Keywords:
Attributive graph
Clustering algorithm
Signal similarity
Heat diffusion
Community detection

*Corresponding Author's Email Address:
Saeedfarzi@kntu.ac.ir

Abstract

Background and Objectives: Simplicity and flexibility constitute the two basic features for graph models which has made them functional models for real life problems. The attributive graphs are too popular among researchers because of their efficiency and functionality. An attributive graph is a graph the nodes and edges of which can be attributive. Nodes and edges as structural dimension and their attributes as contextual dimension made graphs more flexible in modeling real problems.
Methods: In this study, a new clustering algorithm is proposed based on K-Medoid which focuses on graph's structure dimension, through heat diffusion algorithm and contextual dimension through weighted Jaccard coefficient in a simultaneous matter. The calculated clusters through the proposed algorithm are of denser and nodes with more similar attributes.
Results: DBLP and PBLOG real data sets are applied to evaluate and compare this algorithm with new and well-known cluster algorithms.
Conclusion: Results indicate the outperformers of this algorithm in relation to its counterparts as to structure quality, cluster contextual and time complexity criteria.

Introduction

The simple graphs are consisting of nodes and edges collectives which indicate the graph's structural dimension. The simple graphs are applied widely in modeling things and with different counter dependencies, like as friendship and kinship in verity of efficiency realms like analyzing social network, wide web networks and sensor networks. Though, attributive graphs are simple graphs where nodes and edges can have particular features; these features indicate the graph's contextual dimension. Attributive graphs are applied as basic models in running assessments in human interactions in social systems. The graph's structural feature is indicative of individuals and communities in social science while the contextual features is indicative the individuals features' and communities thereof determining social distinction is contributive in constructing a functional graphs evaluation [1]. Today, due to public functional applications, like indicating important modules in biological graph [2] [3][4], data collection related to web pages [5] and determining events/orientations in social networks [6] [7] are of major concern. Traditional algorithms are using structural features to identify communities. The structural features, exhibited through attributive nodes and edges, are beneficial for compact connected components' distinction, like communities. However, in real world networks, node's/edge's features

are too important in studying contextual function's and network's evolution. The attributive graph is an extended graph where in studying in attributive nodes and edges are applied. The node's features are indicative of particular attributions for node's contextual and semantic descriptions. In a similar sense, an edge's features introduce the type of the connection among them. An attributive graph presents a partial model reached in real world network instead of a traditional one [8]. In today's practice, the traditional clustering graph methods are being applied in attributive graphs clustering. The focus of methods is on either topological structural or on graph contextual features where each one of the clusters containing homogenous nodes with respect to the nodes and edges features, while, many recent methods apply a combination of structural and contextual information in attributive graphs clustering [9][10][11].

Considering the definition of attributive graph, the attributive graphs' clustering algorithm, as a compact connected nodes classification, is defined through homogenous features volume. The most important challenge here is to find the harmony between structural and contextual similarities of clusters' nodes and edges [12]. In this study, a new clustering algorithm is proposed for weighted attributive graphs. This algorithm is based on K-Medoid clustering algorithm where both the structural and contextual graph information are applied. A balancing factor is applied in order to balance the structural and contextual data on clustering results. The main idea is transmission of weighted attributive graphs into spatial space where the structural and contextual similarities among nodes are unified. In this proposed algorithm, a similarity criterion, based of heat diffusion [13] is employed with the objective of structural information integration, and, Jaccard weighted similarity criterion [14] is applied for contextual information integration, thus the points in new spatial space would include unified structural and contextual information. Now, K-Medoid algorithm proposed for clustering in spatial space can be applied in this space. The level of structural and contextual features' contribution can be regulated through the balancing factor [16]. Contrary to [11], clusters count and initial main seeds constitute the two basic parameters which should be determined before clustering algorithm is applied. The available algorithms apply nodes degree to find initial seeds in K-Medoid.

Here, clusters count is determined by user and is known as one of the parameters for this proposed algorithm, while the initial seeds are determined based of their degree centrality. Here, the initial seeds selection takes shorter time. Eventually, the proposed algorithm optimizes an objective function which maximizes the inner cluster similarities and minimizes the intra cluster similarities.

The two real data sets, PBLOG [10] including 1490 nodes and DBLP [10] including 10000 nodes, are utilized for experimental evaluation. Many evaluation scenarios are followed to analyze this proposed algorithm. In addition to assessing the algorithm parameters, this algorithm is compared with five other advanced: S-cluster [9], W-cluster [9], SA-cluster [9], SI-Cluster, and KSNAP [16] algorithms based on density and entropy criteria. The results here indicate the outperformers of this algorithm on its counterparts, although in some cases the obtained results are comparable. As to running time complexity, this algorithm outperforms its counter parts as well.

The rest of this paper structure is as follows: literature review is presented in Sec.2; the issues are explained Sec.3; the method is described in Sec.4; empirical studies are presented in Sec.5 and the article concluded in Sec.6.

Literature Review

Most of the graphs clustering algorithms focus on structural aspects based of different objective functions [17][30], as normalized cuts [18] and overall density [19]. The outputs of these algorithms are clusters of high density, but in these algorithms the node features of are ignored.

Today, the graph clustering algorithms focus mostly on structural and contextual aspects of an attributive graph [28][29]. Metis and Markov clustering applied CODICIL [20] clustering for content combination and similar links. The possibility of an edge belonging to one cluster is applied in estimating link similarity, while Jaccard coefficient is applied in estimating contextual similarity.

SI-Cluster [11] is a clustering algorithm based on signal similarity which introduces weighted Jaccard similarity for combining structural and contextual similarities.

Similar to SI-Cluster, this algorithm applies a balancing factor to establish equilibrium among the structural and contextual attributes. Contrary to SI-Cluster method, this applies signal similarity transmission to provide structural information.

Heat diffusion is utilized in this algorithm. Here a simpler and faster method is applied in order to find the initial main seeds, while in SI-Cluster algorithm a complex and time consuming method is employed. SA-cluster [9] is a random walk based clustering method which introduces distances unification for structural and contextual integration. The given graph is clustered based on specific count of clusters. In comparison with SA-cluster, S-cluster is introduced by [19], is of higher structural similarities and lower contextual similarities. The focus of KNSAP algorithm [16] is on contextual

aspect which accumulates nodes of similar features in one cluster.

The Issue

An attributive weighted directed graph $G(V,E,C,A,W)$, where $V = \{v_1. v_2. \ldots . v_{|v|}\}$ is the set of nodes, $E = \{\langle v_i. v_j\rangle | v_i. v_j \epsilon V\}$ is the set of directed edges with weighted function of $C_{ij} \epsilon C$ and $P = \{p_1. p_2. \ldots . p_{|p|}\}$ is a $|p|$ number of features for nodes' description. p_q feature is related to a v_i node, $p_q(v_i) \epsilon P$, with domain of $d_q = |Dom(p_q)|$ vector attached to the vi through $Wq(v_i) \in W$ weight. Thus node v_i has a feature vector $\vec{P}(v_i)$ with $|\vec{P}(v_i)| = \sum_{i=1}^{m} d_i$.

A directed weighted collaborative graph where the nodes introduce the writers' and the edges introduce the colleagues of an article is drawn in Fig. 1. For every writer, 2 adjectives of research interest and mother language are of concern.

Clustering of a directed weighted attributive graph is partitioning the given graph into k subgraph $G^k = (V^k, E^k, C^k, A^k, W^k)$ where $V = U_{k=1}^{K} V^k$, $E = U_{k=1}^{K} E^k$, $V^i \cap V^j = \emptyset$ for any $i \neq j$.

Finding an appropriate equilibrium for optimizing two independent objectives or even a contradictory one is subject to: 1) Structural objective: the inner cluster nodes are similar structurally and different among clusters and 2) Contextual objective: the inner nodes are similar contextually and different among clusters.

In order to balance these two objectives, a balancing factor, $\lambda \epsilon [0.1]$, combines both functions linearly.

$$O_f = \lambda \times O_{str} + (1 - \lambda) \times O_{con} \quad (1)$$

where, Ostr and Ocon are structural and contextual objective, respectively.

Two important questions must be answered in order to design a clustering algorithm:

1) How are the structural and contextual similarities calculated?

2) How is the objective function optimized?

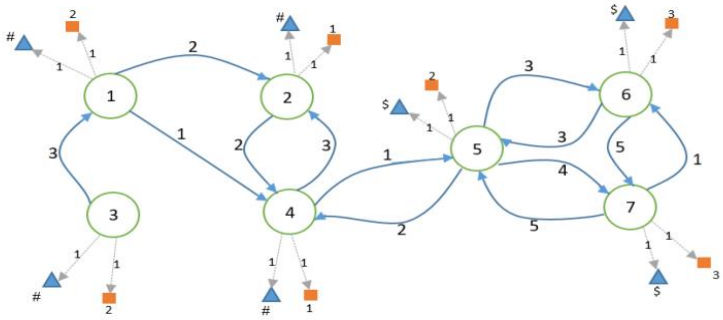


Fig. 1: An example for a directed weighted collaborative graph.

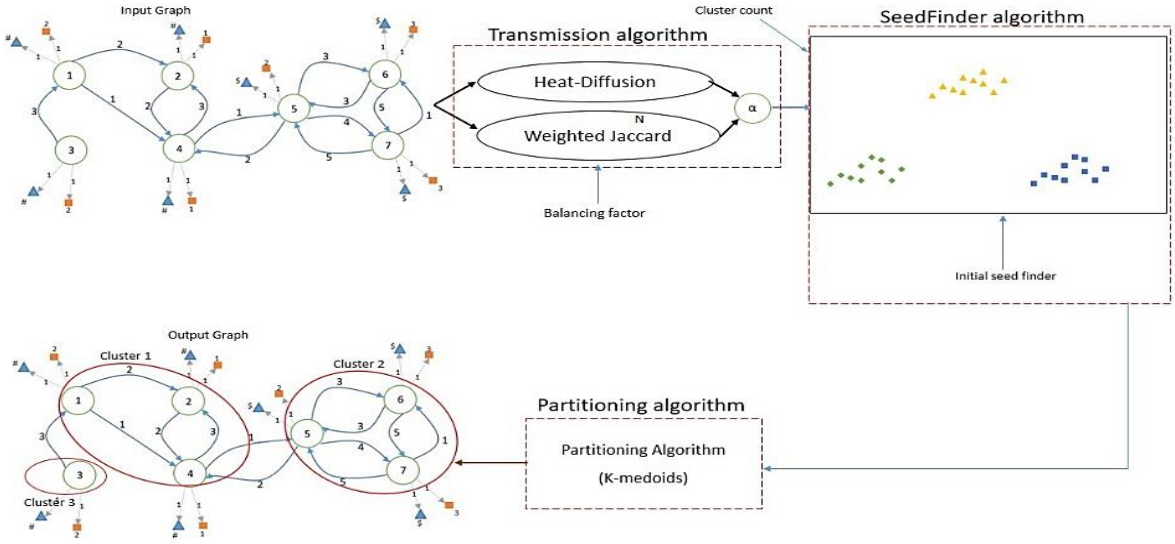


Fig. 2: System's architecture.

To answer the first question, the structural similarities are calculated through heat diffusion and contextual similarities are calculated through weighted Jaccard coefficient. Balancing factor is applied in order to integrate the obtained information and provide a new space. Points in this new space include graph's structural and contextual information. As to answer the second question, due to simplicity and low implementation time algorithm, K-Medoid algorithm is applied in this algorithm to optimize Objective function. The clusters and initial seeds count are the two main challenges which must be of concern. Cluster count is determined by the user, while for key initial seeds selection, the nodes with higher degree would be selected as initial seeds.

The Proposed method

As shown in Fig. 2, the proposed algorithm, here after H-cluster is considered a directed weighted attributive graph as an input and a partitioned graph as output. H-cluster consists of three main parts: 1) Transition algorithm 2) the initial seeds finder algorithm and 3) partitioning algorithm. Transition algorithm, converts the structural and contextual information into spatial space. To do so a heat diffusion similarity [13] and a weighted Jaccard similarity [14] are applied in measure the structural and contextual information. The initial seeds are the main challenge for clustering algorithm [21]. To come up with this problem, the seed-finder algorithm is applied in the second part of the algorithm. Partitioning algorithm finds the optimized clusters through initial main seeds finding and determines their count. The objective function is a defined through the combined structural and contextual similarities. Transition algorithm, seed-finder and partitioning algorithm will be discussed further.

A. Transition Algorithm

In this algorithm, first, the graph's structural data is calculated through heat diffusion simulation, and the contextual information graph is calculated through weighted Jaccard coefficient. At the end a linear combination of these two criteria is calculated through balancing factor.

- Structural transition

Structural transition is a graph model based on heat diffusion [13]. This model can be implemented on both directed and un-directed graphs. Heat diffusion is a physical phenomenon. In an environment, heat always flows from a point with higher temperature to a point with lower temperature. Recently, heat penetration based methods are applied in different aspects like as classification and dimensions reduction [22] [23] [24]. In this article, heat diffusion is applied for modeling the structural similarity in attributive graphs.

To transfer structural and topological data presented through a given attributive graph to spatial space, heat diffusion [13], which is a popular algorithm for data transition, is applied with the objective of calculating similarities between two nodes. According to [25], heat diffusion similarity are more efficient compared to other likewise similarity functions such as Jaccard and Cosine. Heat diffusion is main base for heat diffusion similarity calculation. To calculate the heat diffusion similarity for a graph with n number of nodes, every node is considered as a heat source. During an iterative process, each node is selected as an initial heat source to stimulate all other nodes in the given graph. The process begins with attributing one heat unit to node. After heat diffusion, the initial node and all its neighbors, store the heat in a vector of n number of dimension and re-transfer it to all neighbors. After one step (f(1) time), the source's nodes effect on the whole chart with and the received heat volume in the n dimension vector of nodes is calculated. The mentioned heat diffusion method can be described as a simple and clear mathematical process according to (2)- (9)

$$\frac{f(t+\Delta t)}{\Delta t} = \alpha(H-D)f(t) \tag{2}$$

where

$$H_{ij} = \begin{cases} 1 \ (v_i, v_j) \epsilon \ E \text{ or } (v_j, v_i) \epsilon \ E \\ 0 \ i = j \\ 0 \text{ otherwise} \end{cases} \tag{3}$$

and

$$D_{ij} = \begin{cases} d(v_i) & i = j \\ 0 \text{ otherwise} \end{cases} \tag{4}$$

where, d(vi) is the degree of node vi. The D matrix is a diagonal matrix.

Thus, better exhibition all D and H matrix inputs are normalized based of each nodes degree. D and H matrices can be normalized through the following equations:

$$H_{ij} = \begin{cases} \frac{1}{d(v_i)}, (v_i, v_j) \epsilon \ E \\ 0 \ i = j \\ 0 \text{ otherwise} \end{cases} \tag{5}$$

and

$$D_{ij} = \begin{cases} 1 \ i = j \\ 0 \text{ otherwise} \end{cases} \tag{6}$$

The following differential equation is applied to solve this problem:

$$\frac{d}{dt} f(t) = \alpha t(H-D)f(t) \tag{7}$$

To solve this issue we have:

$$\hat{S} = e^{a(H-D)} f(1) \tag{8}$$

where, d(v) is the node's v degree, and $e^{a(H-D)}$must be yielded through Eq.9:

$$e^{a(H-D)} = I + \alpha(H-D) + \frac{\alpha^2}{2!}(H-D)^2 + \frac{\alpha^3}{3!}(H-D)^3 + \cdots \quad (9)$$

Eventually, the normalized matrix $\hat{S}$ includes heat diffusion data transmitted among different nodes.

The $e^{a(H-D)}$ matrix is named the diffusion core, which is repetitive heat diffusion after initial diffusion.

- Contextual transition

The Jaccard similarity is a common function applied widely on calculating the similarities among two sets. Due to the weighted attributives of the considered graph, the weighted Jaccard similarity function [26] is applied. Thus, in order to convert the contextual information, through attributive graphs, into the spatial space of n dimensions, the weighted Jaccard weighted similarity [26] is calculated between two vectors of attributive nods through (10):

$$\hat{C}_{ij} = \frac{\sum_{q=1}^{\sum_{i=1}^{m} d_i} \min\left(\vec{P}_q(i), \vec{P}_q(j)\right)}{\sum_{q=1}^{\sum_{i=1}^{m} d_i} \max\left(\vec{P}_q(i), \vec{P}_q(j)\right)} \quad (10)$$

where $\vec{P}_q(i)$ and $\vec{P}_q(j)$ are the non-negative features vectors of nodes i and j, respectively and Max(…) and min(…) are the maximum and minimum functions, respectively.

- The proposed transition algorithm

The transition algorithm is adopted to integrate a spatial space through contextual and structural data combination. To make this spatial space, a combination of normalized linear structural and contextual similarity is applied as follows:

$$N_{ij} = \lambda \times \hat{S}_{ij} + (1-\lambda) \times \hat{C}_{ij} \quad (11)$$

here, Nij calculates the contextual and structural similarities between i and j nodes. The transition algorithm is exhibited in Fig. 3.

Max(…) and Min(…) are the maximum and minimum functions, respectively, described in (10).

B. Parameter Determination

High degree nodes in a cluster play most important roles. As to they usually are considered as headers these nodes are surrounded by nodes with lower degree [11]. Thus, in order to find the cluster's initial seeds, first, the nodes must be sorted in descending degree and then, αk numbers of nodes with higher degree are selected, where k is the number of clusters and α is per-defined constant value. The first node is selected as the initial seed. The next node will be selected when it keeps the greatest distance from previously selected nodes as seeds. This process continues until k initial seeds is determined, the time complexity of this algorithm is $O(|V|\log|V| + |V|k) \cong O(|V|\log|V|)$.

Input *A: Attribute vector of* $\sum_{i=1}^{m} d_i$
λ: Balancing factor
t: Total steps
*P: Adjacency Matrix of n*n*

Output *N: Matrix of n*n*

1: ***function*** *Transition_Algorithm*
2: ***begin***
3: *S, C , N : Matrix of n*n*
4: *I: identity Matrix of n*n*
5: $S = e^{a(H-D)} f(0)$
6: ***for*** *i=1 to n*
7: ***begin***
8: $C_{ij} = \frac{Min(i,j)}{Max(i,j)}$
9: $N_{ij} = \lambda \times S_{ij} + (1-\lambda) \times C_{ij}$
10: ***end***
11: ***end***

Fig. 3: Transition algorithm.

Though, the nodes with higher degree are more important and most probably are being as seed, selecting two nodes with higher degree next to each other is not a good idea. Here, the node with higher degree is selected as a seed and the nodes with lower degree are ignored. Seed-finder algorithm implementation is as follows:

Dist(…) represents the distance between the candidate node (Pri) and node j which is selected as a seed previously and S(…) is the matrix of total distances between candidate nodes and all that were selected and confirmed seeds.

C. Partitioning algorithm

The objective is to bring the inner cluster similarity to maximum and minimize the inter-cluster similarity. In order to propose a compatible partitioning algorithm, two issues must be of concern: 1) the manner of contributing each of the nodes to a cluster and 2) the manner of defining an intra-cluster objective function for having access to intra-clusters higher structural and contextual similarities and lower enter-cluster structural and contextual similarities. As to issue one, a K-Medoid based compatible partitioning algorithm is introduced [27], which is an extended version of K-Means clustering algorithm where the central seeds in a cluster are selected by Medoids. In K-Medoid based partitioning algorithm, Medoids are selected through the initial key seeds, which are identified by seed finder algorithm. During each iteration, the seed are determined once more and the points are reallocating to the clusters. The process continues while the seeds stop changing in next iteration.

```
Input: A // adjacency matrix
Output: Seeds // highest central nodes
1: function SeedFinder
2:     Seeds: Set =[]
3: begin
4:     Pr: Array // Pr.size= A.cols
5:     for each col of A
6:      begin
7:        Pr_i = Σ_j A_ij / A.cols + Σ_j A_ij / A.rows
8:      end
9:     Pr= Sort(Pr)
10:    Seeds.add(S_1 ∈ Pr)
11:    while Seeds.size==k
12:    begin
13:       for i=2 to Pr.size
14:       Begin
15:          for j=1 to Seeds.size
16:          begin
17:            S(i,j)= S(i,j) + Dist(Pr(i), Seed(j));
18:          end
19:       end
20:       Seeds.add(mas(S))
21:    end
22: end
```

Fig. 4: The initial seed finder algorithm.

As to second issue, the density and entropy volume are applied, respectively, to calculate the structural and contextual similarities. Thus, the objective function is calculated through (2):

$$O_f = \lambda \times O_{str} + (1-\lambda) \times O_{con} = \lambda\, D\left([G^j]_{j=1}^{j=k}\right) + (1-\lambda) * 1/E([G^j]_{j=1}^{j=k}) \quad (12)$$

where, λ represents the balance factor D(...) is a density function applied in structural similarity calculation and the contextual similarity is calculated through E(...), the entropy function.

The density function reveals the intensity of coherency of clusters' compression. The partitioning graph's overall density is yielded through (13):

$$D\left([G^j(V^j,E^j)]_{j=1}^{j=k}\right) = \sum_{j=1}^{k} \frac{\sum_{\langle v_p,v_q\rangle \in E^j} C(\langle v_p,v_q\rangle)}{\sum_{\langle v_p,v_q\rangle \in E} C(\langle v_p,v_q\rangle)} = \frac{1}{\sum_{\langle v_p,v_q\rangle \in E} C(\langle v_p,v_q\rangle)} \times \sum_{j=1}^{k} \sum_{\langle v_p,v_q\rangle \in E^j} C(\langle v_p,v_q\rangle) \quad (13)$$

where, C(...) is the graph's cost function.

Also the total cluster entropy is applied to compute relevancy of nodes of a given graph with respect to attribute values, Eq. 14.

$$E\left([G^j(V^j,E^j,A^j)]_{j=1}^{j=k}\right) = \frac{1}{k}\sum_{j=1}^{k} E^c(G^j(V^j,E^j,A^j)) \quad (14)$$

$$E^c(G^j(V^j,E^j,A^j)) = -\frac{1}{m}\sum_{q=1}^{m}\sum_{a_q \in d_q} P(a_q,V^j)\log\, P(a_q,V^j) \quad (15)$$

$$P(a_q,V^j) = \frac{|v_h \in v^j|\, a_q(h) = a_q|}{|v^j|} \quad (16)$$

The more the placement of nodes with similar features in a cluster, the lowers the entropy.

Empirical studies

Functionality of this method is evaluated through many experiments. The three main scenarios consist of: 1) assessing the convergence of the proposed algorithm, 2) assessing the parameters' effect on cluster's quality, and running time complexity and 3) analyzing the density and entropy of this method compared to other well-known clustering algorithms.

A. Data

Two real data sets called (PBIOG) [10] and (DBLP) [10] are applied in order to evaluating the efficiency of introduced method, Table 1.

Table1. Several statistics related to data collections

	PBLOG	DBLP
Description	Political Weblog network	Co-authorship network
Nodes	1490	10000
Edges	33433	65734
Attibutes (values)	Political leaning	Prolific Primary topic

B. Configuration

This experiment is run on Intel® coreTM with seven 2.80 GHz cores and 6 G main RAM. The code is implemented MATLAB. This H-cluster method is compared with five contemporary algorithms of:, SA-Cluster, S-Cluster, W-Cluster, SI-Cluster[1]and KNSAP.The SA-Cluster algorithm is introduced by [9], where both the structural and contextual aspects of the network are of concern. The nodes closeness is calculated through random walk S-Cluster algorithm [9] where only the structural aspect is of concern. W-Cluster is presented by [9]. where the both the structural and contextual aspects are combined through a linear combination [9]. KNSAP is introduced by [16] and SI-Cluster is proposed by [11], where though both, the aspects, are combined through a

[1]SI-Cluster using signal diffusion without Seed Finder algorithm like [11]

linear combination accumulated the nodes with similar features in a cluster. To compare the function and efficiency of these algorithms the density and entropy criterion are calculated through (14) and (15).

C. Convergence

The objective function of every iteration of H-Cluster with or witout seed finder algorithm on PBLOG and DBLP data collection is shown in Fig. 5. The H-Cluster without seedfinder applies random function to find the seeds. The quantities are expressed as follows: (Ma teration=20 , λ = 0.5,$\alpha = 40$ (k=10 for DBLP and k=9 for PBLOG))

As observed in Fig. 5-a H-Cluster reaches convergence faster than other. As observed in Fig. 5, the maximum objective volume for both DBLP and PBLOG data sets are obtained, at the third iteration that is the initial main seeds determined based on seed finder algorithm provides the means for clustering algorithm to obtain better results in little iteration.

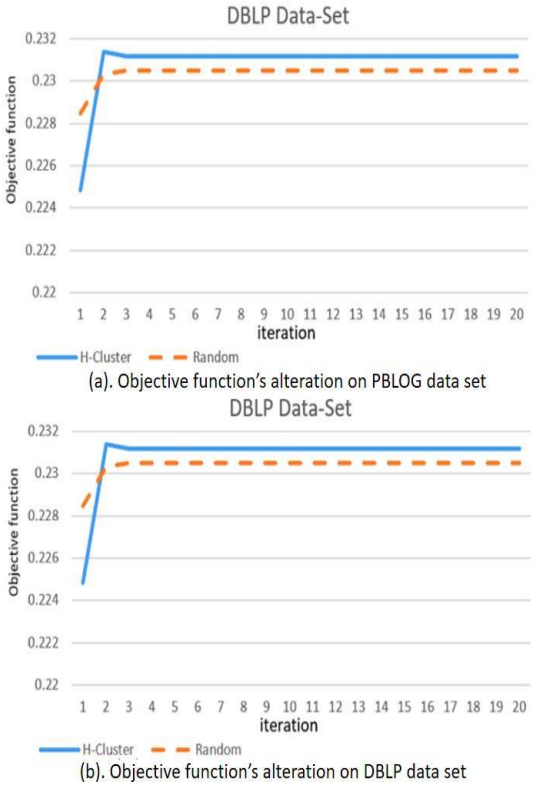


Fig. 5: Objective function alterations at every iteration.

The two main parameters are analyzed as follows: $\alpha\ (alfa)$ which applied in finding initial seeds and λ which is a structural and contextual balancing factor. The λ effect on cluster quality on DBLP data set when $\alpha = 40$ and the number of clusters is 10. Fig. 6.

Fig. 7 and Fig. 8 show Alfa's (α) effect on cluster's quality on (density, entropy and objective function) in PBLOG and DBLP datasets respectively. Here $\lambda = 0.5$ $\mathrm{K(PBLOG)} = 3$ and $\mathrm{K(DBLP)} = 10$.

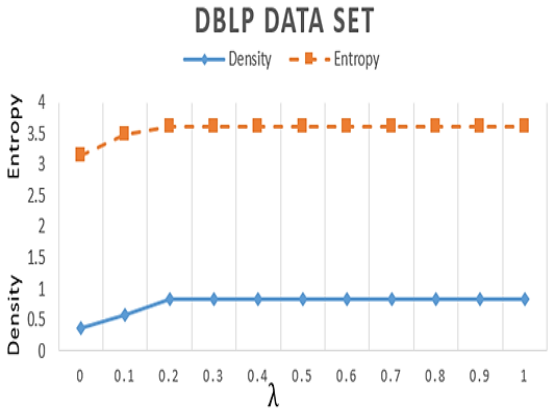


Fig. 6: Balancing factor effect on density and entropy criterion in DBLP data set.

As observed in Fig. 7(a) and Fig. 8(a) the higher the α volume the more the density. The best volume of entropy on PBLOG dataset, $(\alpha) = 5$, Fig. 7(b) and according to in Fig. 8 the best volume of entropy for DBLP dataset is $(\alpha) = 40$. After balancing the density and entropy criterion, as it Fig. 7(c) and Fig. 8, the volume of objective function is obtained. The maximum volume of objective function in PBLOG dataset is $(\alpha) = 5$ and the same volume for DBLP is $(\alpha) = 40$.

D. Running Time

Here, Alfa (α) is the most effective factor in time complexity. The H-Cluster running time with respect to Alfa's (α) alternations in DBLP data sets illustrated in Fig. 9.

As observed in Fig. 9, an increase α volume, increases the algorithm's running time, leading to an important in cluster's quality Fig. 10.

E. Cluster Algorithms Comparisons

To compare H-cluster algorithm with other new clustering algorithms, several evaluations are run on PBLOG and DBLP datasets. Here, λ=0.5 for S-Cluster, SA-Cluster, SI-Cluster and H-Cluster and α=5k for H- cluster, SI-Cluster on PBLOG dataset with α=40k for these algorithms for DBLP dataset. The laboratory experiments' results on density and entropy in PBLOG dataset are expressed in Fig. 10.

In average, SI-Cluster (-0.01042 | +0.0887 | +3.6060) and (+0.21356 | +0.08412 | +14.549) outperforms SA-Cluster and H-Cluster as to (density | entropy | objective function) criteria, respectively. In average, H-Cluster, as to the (density | entropy | objective function) (-0.2240 | +0.00458 | -10.9431) (-0.224 | -0.08412 | -14.4590) fails compered to SA-Cluster and SI-Cluster, respectively. As observed in Fig. 10, considering the fact that both structural and contextual aspects are of concern in H-Cluster, indicating that H-Cluster as to PBLOG dataset is not as efficient as SI-Cluster, the SI-Cluster outperforms H-cluster on PBLOG dataset.

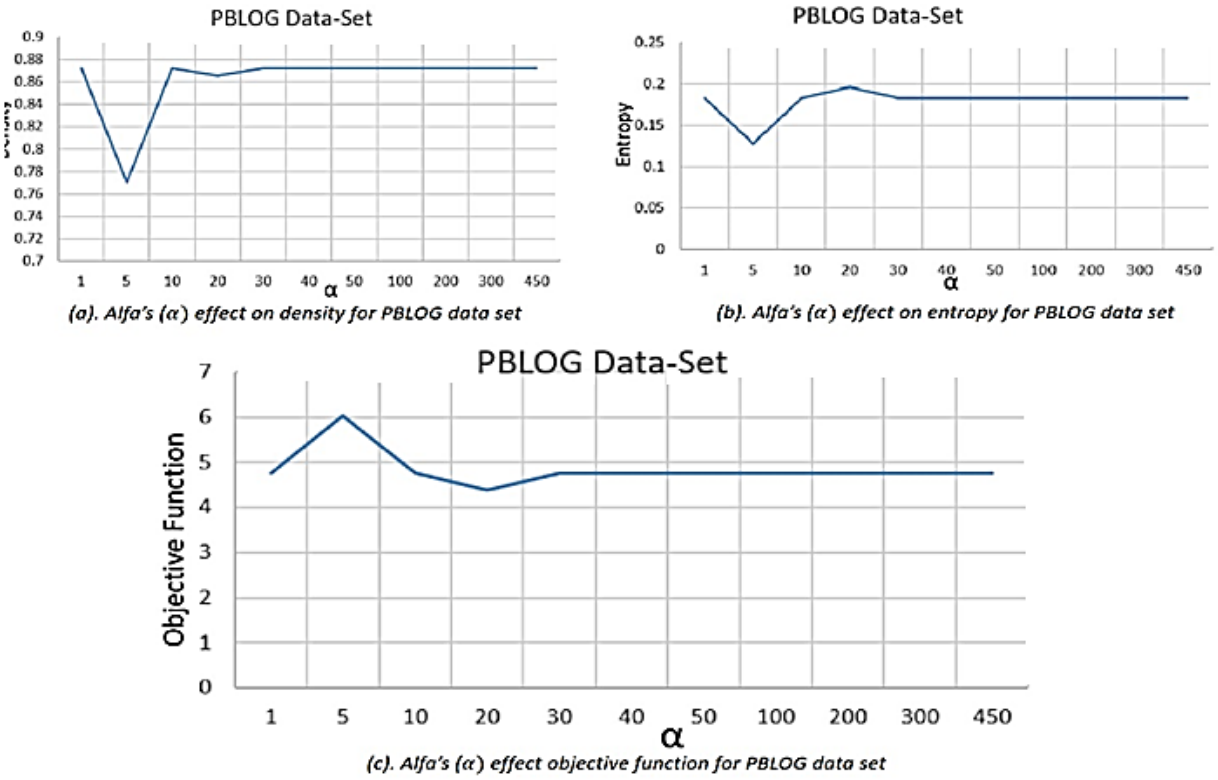


Fig. 7: Alfa's (α) effect on density and entropy criterion and objective function of (PBLOG).

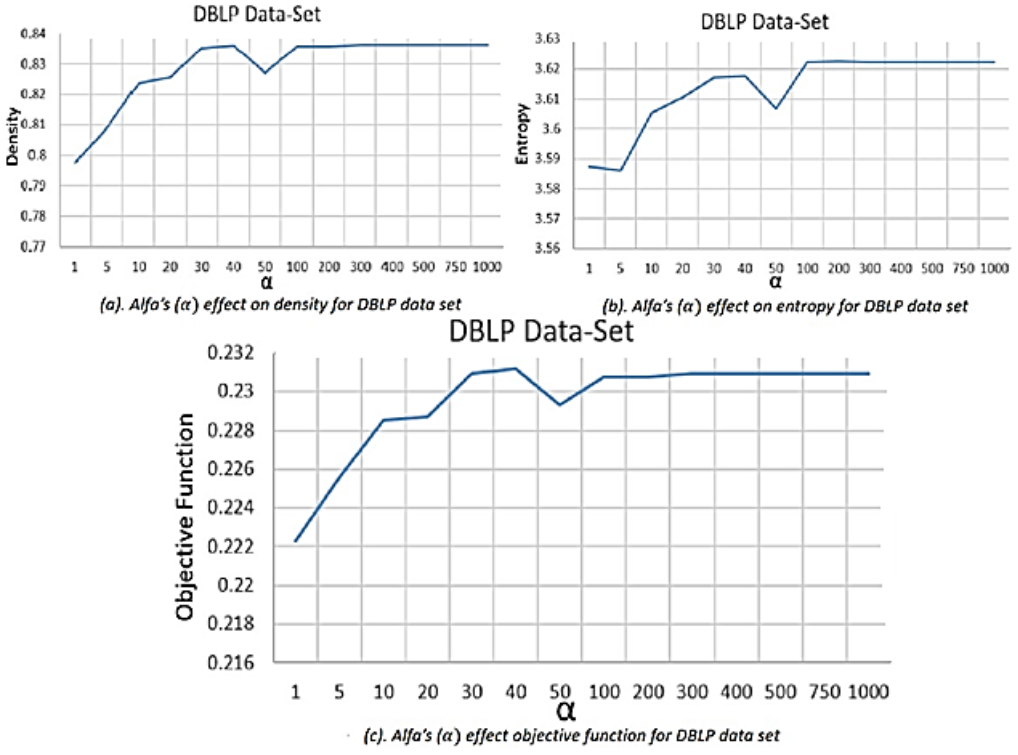


Fig. 8: Alfa's (α) effect on density and entropy criterion and objective function of (DBLP).

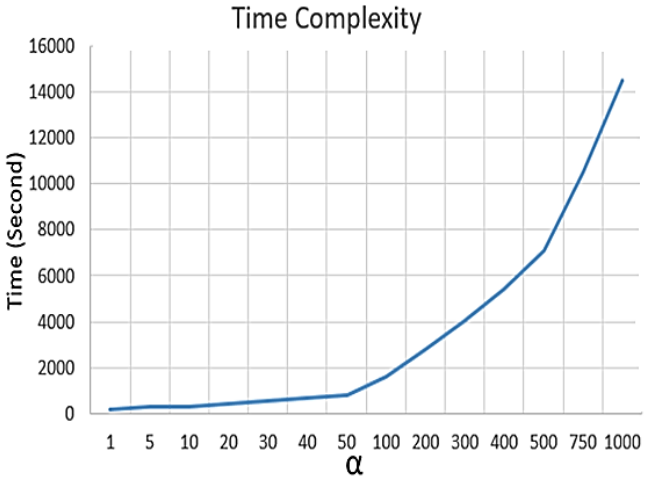


Fig. 9: Alfa's (α) effect on DBLP dataset time complexity.

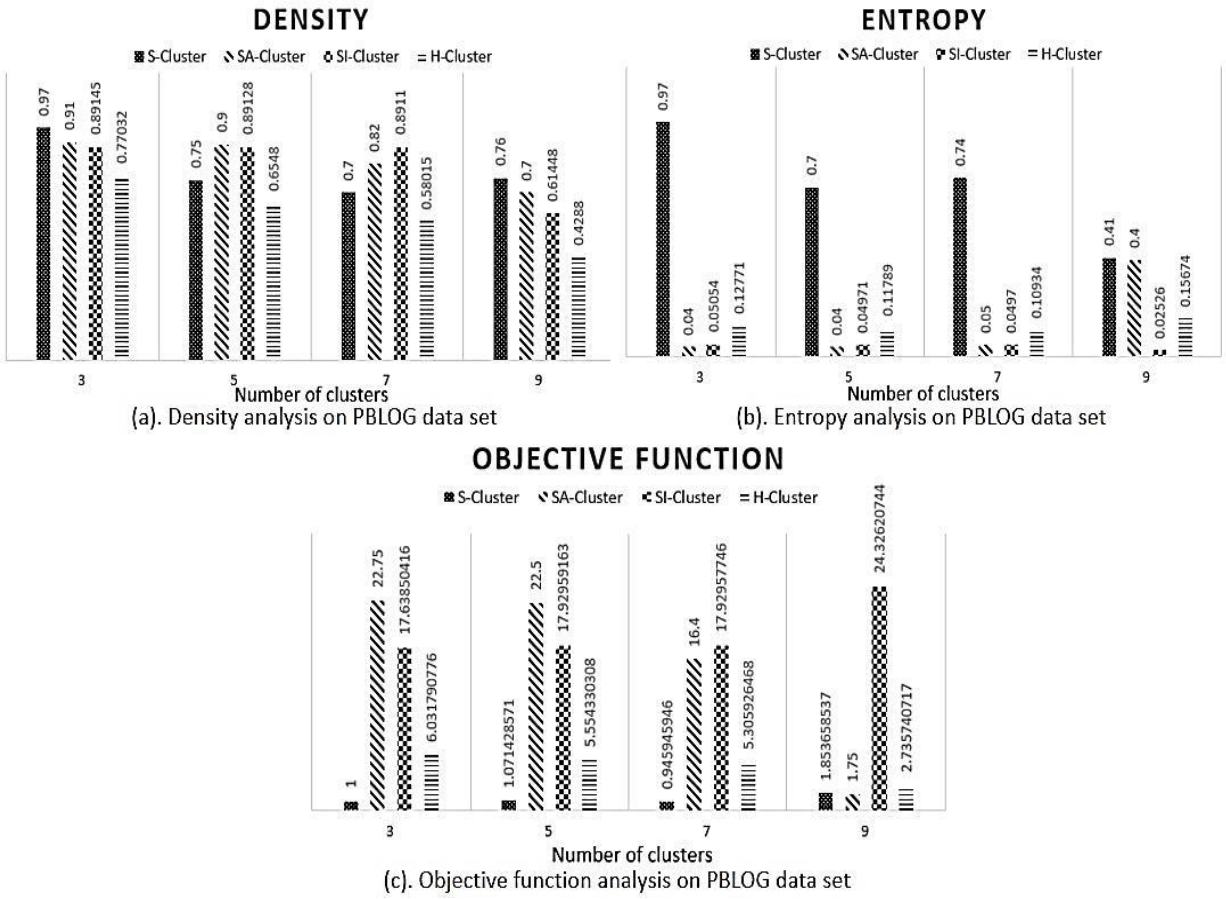


Fig. 10: Density, entropy and objective function's analysis on *PBLOG data set* ($\lambda = 0.5$, $\alpha = 5$ k for PBLOG).

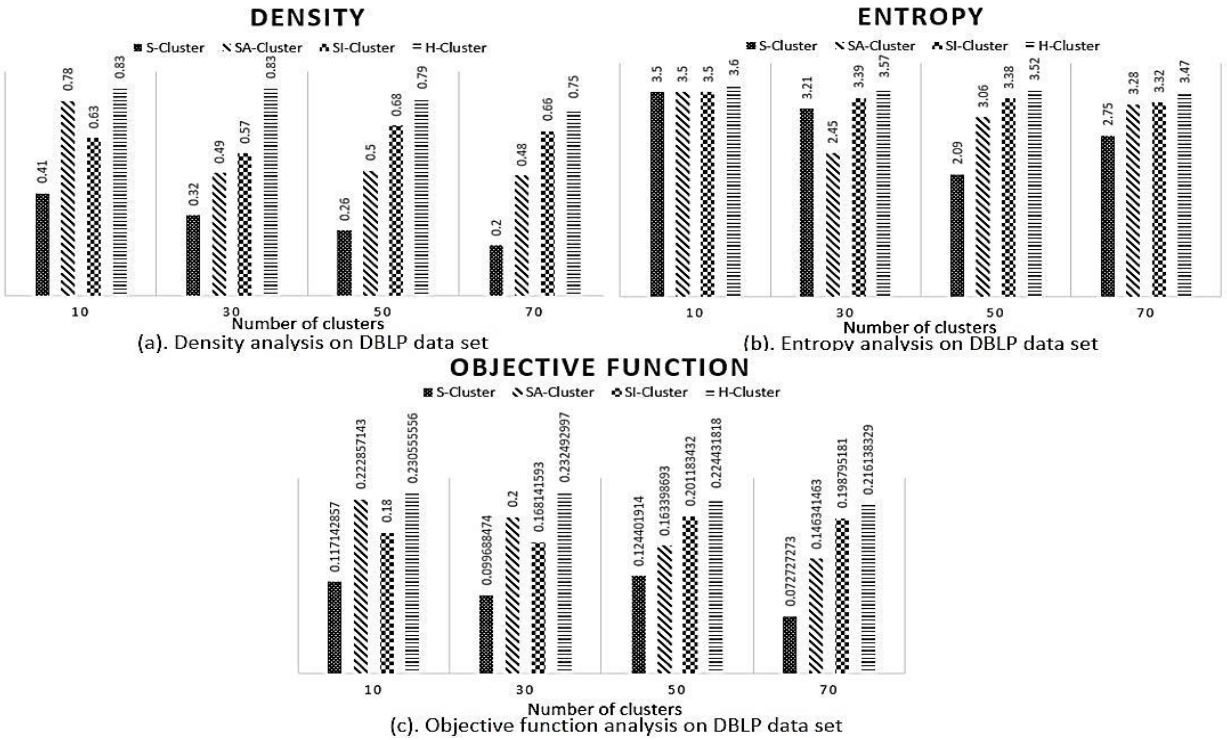


Fig. 11: Density, entropy and objective function's analysis on DBLP data set ($\lambda = 0.5$ $\alpha = 40$ k for DBLP).

In average, H-Cluster, as to the (density | entropy | objective function) criteria measurements (-0.1865 | +0.57708 | +3.68919) are improved compared to S-Cluster. S-Cluster algorithm considers only structural aspect, thus, as observed, H-Cluster outperforms S-Cluster as to entropy and objective function criteria.

In general, the obtained results indicate the outperformers of SI-Cluster advantages compared to other known algorithms as to density, entropy and objective function measurement on PBLOG dataset.

The laboratory experiments' results on density and entropy in DBLP dataset are expressed in Fig. 11.

In average, H-Cluster (+0.2375 | -0.4675 | +0.042755) (+0.165 | -0.1425 | +0.038875) outperforms SA-Cluster and SI-Cluster as to (density | entropy | objective function) criteria, respectively. As observed in Fig. 11, though in H-Cluster algorithm considers both the structural and contextual aspects are of concern, it is observed that this algorithm generates more density clusters compared to that of SA-Cluster in all aspects, in addition to, H-Cluster generating more density compared to that of the SI-Cluster algorithm. The SA-Cluster algorithm, similar to H-Cluster and SI-Cluster algorithms, considers both structural and contextual aspects, indicating that H-Cluster outperforms SI-Cluster and SA-Cluster in DBLP dataset. In average, H-Cluster yielded improved (+0.5025 | -0.6525 | +0.122415) on (density | entropy | objective function) when measurements are compared. Though S-Cluster considers only the structural aspect, and, as it observed, this proposed algorithm generates more density clusters for all the issues. In average, SI-Cluster, as to the (density | entropy | objective function) criteria measurements (+0.0725 | -0.325 | +0.003881) (-0.165 | +0.1425 | -0.038875) fails compered to SA-Cluster and H-Cluster, respectively. Except in entropy criteria, failed to SA-Cluster and SI-cluster compared to H-Cluster on entropy criteria. As observed in Fig. 11, considering the fact that both structural and contextual aspects are of concern in SI-Cluster, indicating that SI-Cluster as to DBLP dataset is not as efficient as H-Cluster, the H-Cluster outperforms SI-cluster on DBLP dataset.

In general, the obtained results indicate the outperformers of H-Cluster advantages compared to other known algorithms as to density and objective function measurement on DBLP dataset. Note that small size of PBLOG dataset is one of the main reasons that the proposed algorithm cannot achieve good results.

Conclusion

In this study, an effective method for graph clustering is proposed with the objective of finding cluster with higher density and homogenized nodes as to their features.

The structural and contextual aspects are first, obtained through the given graphs by applying heat diffusion and weighted Jaccard similarities, and next, are integrated into a spatial space. In this space, this newly proposed algorithm seeks to maximize intra-cluster similarity while seeking to reduce the enter-cluster similarity to its lowest level. The clustering results quality is determined through density and entropy criteria. The introduced method outperforms existing algorithms. The empirical studies express the competitive results with respect to the cluster quality and this proposed method operates in polynomial time and is scalable for medium and large scale networks.

Author Contributions

S. Farzi and S. kianian contributed to the design and implementation of the research, to the analysis of the results and to the writing of the manuscript.

Acknowledgment

We thank the editor and all anonymous reviewers.

Conflict of Interest

The author declares that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancy have been completely observed by the authors.

Abbreviations

PBLOG	Political Weblog network
DBLP	Co-authorship network
KNSAP	k-Node Sammarization attribute pair-wise nodes
H-Cluster	Hierarchical Clustering Algorithm
SI-Clustering	Silhouette index Clustering

References

[1] M.E. Newman, "The structure and function of complex networks," SIAM Review, 45(2):167-256, 2003.

[2] R. Guimera, L.A.N. Amaral, "Functional cartography of complex metabolic networks," Nature, 433: 895, 2005.

[3] E. Ravasz, A.L. Somera, D.A. Mongru, Z.N. Oltvai, A.-L. Barabási, "Hierarchical organization of modularity in metabolic networks," Science, 297: 1551-1555, 2002.

[4] D.M. Wilkinson, B.A. Huberman, "A method for finding communities of related genes," in Proc. the national Academy of sciences, 101: 5241-5248, 2004.

[5] Y. Dourisboure, F. Geraci, M. Pellegrini, "Extraction and classification of dense communities in the web," in Proc. the 16th international conference on World Wide Web: 461-470, 2007.

[6] R. Cazabet, H. Takeda, M. Hamasaki, F. Amblard, "Using dynamic community detection to identify trends in user-generated content," Social Network Analysis and Mining, 2: 361-371, 2012.

[7] K. Konstantinidis, S. Papadopoulos, Y. Kompatsiaris, "Exploring Twitter communication dynamics with evolving community analysis," PeerJ Computer Science, 3: e107, 2017.

[8] C. Bothorel, J.D. Cruz, M. Magnani, B. Micenkova, "Clustering attributed graphs: models, measures and methods," Network Science, 3(3): 408-444, 2015.

[9] H. Cheng, Y. Zhou, J.X. Yu, "Clustering large attributed graphs: A balance between structural and attribute similarities," ACM Transactions on Knowledge Discovery from Data (TKDD), 5(2): 12, 2011.

[10] W. Nawaz, K.-U. Khan, Y.-K. Lee, S. Lee, "Intra graph clustering using collaborative similarity measure," Distributed and Parallel Databases, 33: 583-603, 2015.

[11] S. Farzi, S. Kianian, "A novel clustering algorithm for attributed graphs based on K-medoid algorithm," Journal of Experimental & Theoretical Artificial Intelligence, 30(6): 1-15, 2018.

[12] Z. Xu, Y. Ke, Y. Wang, H. Cheng, J. Cheng, "A model-based approach to attributed graph clustering," in Proc. the 2012 ACM SIGMOD international conference on management of data: 505-516, 2012.

[13] H. Ma, I. King, M.R. Lyu, "Mining web graphs for recommendations," IEEE Transactions on Knowledge and Data Engineering, 24(6): 1051-1064, 2012.

[14] M. Popescu, J. M. Keller, J. A. Mitchell, "Fuzzy measures on the gene ontology for gene product similarity," IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB), 3(3): 263-274, 2006.

[15] Y. Zhou, H. Cheng, J. X. Yu, "Clustering large attributed graphs: An efficient incremental approach," in Proc. 2010 IEEE 10th International Conference on Data Mining (ICDM): 689-698, 2010.

[16] Y. Tian, R.A. Hankins, J.M. Patel, "Efficient aggregation for graph summarization," in Proc. the 2008 ACM SIGMOD international conference on Management of data: 567-580, 2008.

[17] M.E. Newman, M. Girvan, "Finding and evaluating community structure in networks," Physical review E, 69(2): 026113, 2004.

[18] J. Shi, J. Malik, "Normalized cuts and image segmentation," IEEE Transactions on pattern analysis and machine intelligence, 22(8): 888-905, 2000.

[19] X. Xu, N. Yuruk, Z. Feng, T.A. Schweiger, "Scan: a structural clustering algorithm for networks," in Proc. the 13th ACM SIGKDD international conference on Knowledge discovery and data mining: 824-833, 2007.

[20] Y. Ruan, D. Fuhry, S. Parthasarathy, "Efficient community detection in large networks using content and links," in Proceedings of the 22nd international conference on World Wide Web: 1089-1098, 2013.

[21] S. Kianian, M.R. Khayyambashi, N. Movahhedinia, "Semantic community detection using label propagation algorithm," Journal of Information Science, 42(2): 166-178, 2016.

[22] M. Belkin, P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," Neural computation, 15(6): 1373-1396, 2003.

[23] I.K. RISI, "Diffusion kernels on graphs and other discrete input spaces," in Proc. 19th Int. Conf. Machine Learning, 2002.

[24] J. Lafferty, G. Lebanon, "Diffusion kernels on statistical manifolds," Journal of Machine Learning Research, 6(5): 129-163, 2005.

[25] Y. Li, C. Jia, J. Yu, "A parameter-free community detection method based on centrality and dispersion of nodes in complex networks," Physica A: Statistical Mechanics and its Applications, 438: 321-334, 2015.

[26] S. Ioffe, "Improved consistent sampling, weighted minhash and l1 sketching," in Proc. 2010 IEEE 10th International Conference on Data Mining (ICDM): 246-255, 2010.

[27] L. Kaufman, P. Rousseeuw, Clustering by means of medoids: North-Holland, 1987.

[28] M. Seifikar, F. Saeed, M. Barati, "C-Blondel: An efficient louvain-based dynamic community detection algorithm," IEEE Transactions on Computational Social Systems 7(2): 308-318, 2020.

[29] M. Fozuni. Shirjini, , S. Farzi, A. Nikanjam, "MDPCluster: a swarm-based community detection algorithm in large-scale graphs." Computing, 102: 893-922, 2020.

[30] S.F. Mirmousavi, S. Kianian, "Link Prediction using Network Embedding based on Global Similarity." Journal of Electrical and Computer Engineering Innovations (JECEI), 8(1): 97-108, 2019.

Biographies

Sahar Kianian received her B.Sc. degree in computer Engineering (2007) from razi University, also M.Sc. and Ph.D. degrees in computer Engineering from Isfahan University (2010 and 2016, respectively). She is an assistant professor of computer engineering at Shahid Rajaee University. Her research interests are the application of algorithms, machine learning and data science to complex networks, focuses on protein interactions, connections of neurons and relationships among people. Applications include disease prediction, drug discovery, event detection and tracking, recommendation system, web mining and social influence mining.

Saeed Farzi received the Ph.D. degree in computer engineering from Tehran University, Tehran, Iran, in 2016. He joined the Artificial Intelligence Department, K. N. Toosi University of Technology, Tehran, in 2017. His research interests include machine learning, information retrieval, and social network analysis.

Hamed Samak received his Bachelor and MA from K. N. Toosi university of technology. His research interests are the application of algorithms, machine learning and data science to complex networks, focuses on protein interactions, connections of neurons and relationships among people. Applications include disease prediction, drug discovery, event detection and tracking, recommendation system, web mining and social influence mining

How to cite this paper:
S. Kianian, S. Farzi, H. Samak, "A new clustering algorithm for attributive graphs through information diffusion approaches," Journal of Electrical and Computer Engineering Innovations, 8(2): 273-284, 2020.

DOI: 10.22061/JECEI.2020.7190.366

URL: http://jecei.sru.ac.ir/article_1472.html

Journal of Electrical and Computer Engineering Innovations
(JECEI)
Journal homepage: http://www.jecei.sru.ac.ir

Research paper

Real-time Implementation of Sliding Mode Control for Cascaded Doubly Fed Induction Generator in both Islanded and Grid Connected Modes

H. Zahedi, G.R. Arab Markadeh, S. Taghipoor*

Department of Engineering, Shahrekord University, Shahrekord, Iran.

Article Info

Article History:
Received 23 November 2019
Reviewed 02 January 2020
Revised 14 February 2020
Accepted 21 April 2020

Keywords:
Cascaded Doubly Fed Induction Generator (CDFIG)
Sliding Mode Control (SMC)
Grid-connected
Relative Gain Array (RGA)
Power control

*Corresponding Author's Email Address:
arab-gh@eng.sku.ac.ir

Abstract

Background and Objectives: Cascaded doubly fed induction generators (CDFIGs) can directly connected to isolated load or power grid without any brushes which are needed in conventional DFIGs. Output control targets before grid connection of CDFIGs are voltage and frequency control and after that are active and reactive power control. In control aspect, output control of CDFIG is a multi-input multi-output (MIMO) subject. In this paper, Relative Gain Array (RGA) methodology, as a MIMO interaction index, is used to show the degree of relevance between the control inputs and output targets, in both voltage control mode (before grid connection) and active-reactive power control mode (after grid connection). Based on RGA results, conventional PI controllers cannot be used to decouple control of generator outputs in grid connected mode. So, a powerful method based on sliding mode approach is proposed to generate the proper control voltages for output control of CDFIG in both islanded and grid connected mode. Simulation and experimental results using Matlab and TMS320F28335 based prototype of CDFIG are provided to demonstrate the effectiveness and robustness of the proposed method.
Methods: A mathematical method based on RGA matrix is used to evaluate the amount of interactions between output targets and input control variables in CDFIGs in islanded and grid connected mode.
Results: Conventional PI controller is a proper method to control the output voltage of Power Machine (PM) in CDFIG but is not a suitable technique for active and reactive power control in grid-tied mode.
Conclusion: Sliding mode control can be used to decouple control of CDFIGs in both before and after grid connection. As well as, robustness against the wind speed variation and parameters uncertainties is proved via both simulation and experimental tests.

Introduction

Variable speed constant frequency (VSCF) operation is used in many industries that require Doubly Fed Induction Generators (DFIG). In these generators, even if there is a change of speed, they can be directly connected to the constant frequency grid using partially rated converters and rotor current regulation. Therefore, they are useful for use in wind and hydropower systems [1].

However, the structure of these generators, which includes brushes and slip rings, reduces the life of the device due to the need for maintenance. Therefore, it is not quite suitable for use in applications such as aircraft, which require high reliability and long maintenance period. In the CDFIG generator, by using two DFIGs and connecting the windings of the rotors to each other, a

new brushless structure is obtained, which can be a suitable alternative to DFIG [2].

CDFIG works like a single DFIG so it can be connected directly to the network.

By connecting the two DFIG rotors mechanically and electrically, a CDFIG is obtained. DFIG machines are known as power machine and control machine [3].

The electric load or the grid is connected to the stator winding of the power machine (PM), and a power electronic converter supplies the stator winding of the control machine [4], Fig. 1.

Several studies on variable speed constant frequency generating system (VSCF) using CDFIG have been performed in applications such as windmills, hydropower and aircraft power supply [5].

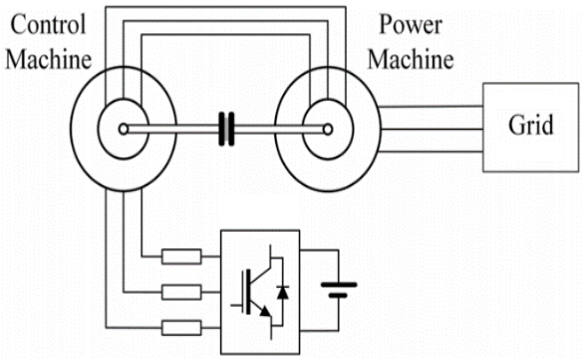


Fig. 1. CDFIG configuration.

The structure of CDFIG is complex in architecture and difficult to control, as well as, de-rating of the machine and higher price in comparison with DFIG. Therefore, despite its advantages, it is not widely used.

Due to the large size of the dynamic equation model, which is obtained by merging two DFIGs, the control of this machine has been influenced by its architecture [6].

A vector control method for DFIG is proposed in [7] and [8]. The method of improving the stability of multi-machine power network (MMPN) has been studied and the application of the conventional non-linear sliding mode control (SMC) is presented in [9].

The terminal SMC as a modified version of the conventional sliding mode control is presented in [10] for DFIG. In [11] by using the sliding mode method and estimating the position of the rotor, the DFIG is controlled.

Efforts to control CDFIG have been limited, mainly based on field oriented control or vector control. In the vector control method, linear controllers such as PI regulate the components of the rotor currents, which include torque (or active power) and rotor flux (or reactive power) [12]-[13].

In [14], the active and reactive power of the power machine is controlled based on vector control using PI controllers. But in this method, all of machine parameters are required in a recursive structure to generate the rotor reference currents and then the control machine reference currents. Also CDFIG is controlled based on DPC for wind energy applications in [15].

In MIMO system, as the interaction between each input with all outputs is high, the controller design will be complicated. As we know, in CDFIG the control inputs are the voltage of CM components and the outputs of plant are the active and reactive powers in grid connected mode or PM voltage components in islanded mode.

In this paper, using a powerful index named as Relative Gain Array (RGA), the decoupling index between the CM voltage components and the generator outputs is calculated. This index describes that the conventional controllers such as Proportional-Integrator (PI) especially in the grid-connected mode are not suitable for control of active and reactive power via PM voltages. Reference [16] used the RGA to calculate the degrees of relevance between the stator voltage and the stator flux of DFIG. The main contributions of this paper can be listed as:

1) The RGA indexes are calculated for output voltage control mode (in islanded mode) and active-reactive control mode (in grid connected mode). Therefore, the complexity of PI coefficients' tuning is clearly stated and proved with RGA index in grid connected mode. In this paper, based on the best knowledge of authors, for the first time, the clear relation between RGA index and complexity of control for CDFIG is reported.

2) In this paper, a sliding mode control is proposed to control the output active and reactive power of CDFIG. The required control voltage of the stator of control machine is calculated directly by the SMC method.

3) In the SMC method, the degree of controller system is decreased from order 6 to order 2, and using a proper Lyapunov function, the control voltage is obtained which is robust to system parameters. The robustness of SMC against the rotor resistance variation and rotor speed due to variable wind speed is proved.

4) The comparison of the proposed SMC and VC is done by experimental tests using a 370w CDFIG machine.

Simulation and experimental results are presented to show the performance and robustness of the suggested control configuration during variations of operating point.

CDFIG Model

The two DFIMs have pole pair p_1 and p_2 respectively, with rotors connected in inverse coupling sequence. Then the voltages and currents of rotors can be written as:

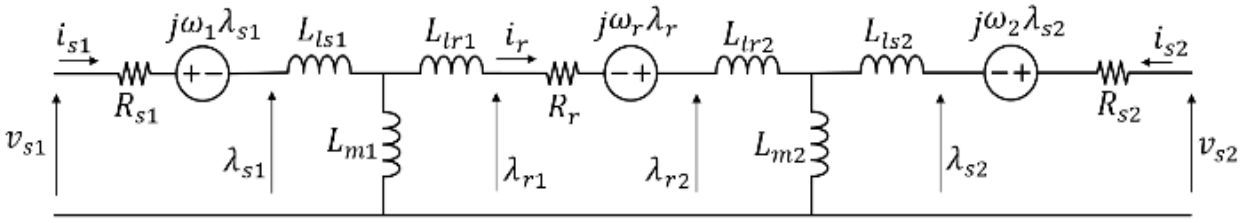


Fig. 2. Equivalent circuit of CDFIG.

$$
\begin{aligned}
v_{qr} &= v_{qr1} = v_{qr2} \\
v_{dr} &= v_{dr1} = -v_{dr2} \\
i_{qr} &= i_{qr1} = -i_{qr2} \\
i_{dr} &= i_{dr1} = i_{dr2}
\end{aligned} \tag{1}
$$

where v_{qr}, v_{dr}, i_{qr} and i_{dr} are the rotor voltage and rotor current components, respectively.

In order to specify the parameters of the two DFIM of the cascade, 1 and 2 subscriptions are used for all values in the system.

The subscripts s or r show the quantities of stator or rotor, and number 1 or 2 show the quantities of first or second machine, and q or d show the quantities of q-component or d-components.

Due to the d-components of stator voltage and q-component of rotor currents of control machine and power machine are in opposite sign.

The rotor is rotated at the mechanical angular frequency ω_m and the frequency of PM voltage is ω_1. So, the flux that induced into the rotor bars has a frequency equal slip frequency of power machine that shown by (ω_r),

$$\omega_r = \omega_1 - p_1 \cdot \omega_m \tag{2}$$

Because the phase sequence is reversed at the rotor connection point, the flux wave frequency caused by the rotor control machine with the power machine slip frequency is in the opposite direction,

$$\omega_2 = -(\omega_1 - (p_1 + p_2) \cdot \omega_m) \tag{3}$$

In the other words, ω_m can be shown as

$$\omega_m = \frac{\omega_1 + \omega_2}{p_1 + p_2} \tag{4}$$

The behavior of CDFIG can be described by equations for each stator and combination of rotors. The complete CDFIG dynamic model in d-q reference frame can be given as:

$$
\begin{aligned}
v_{qs1} &= R_{s1} \cdot i_{qs1} + \frac{d}{dt}\lambda_{qs1} + \omega_1\lambda_{ds1} \\
v_{ds1} &= R_{s1} \cdot i_{ds1} + \frac{d}{dt}\lambda_{ds1} - \omega_1\lambda_{qs1} \\
v_{qs2} &= R_{s2} \cdot i_{qs2} + \frac{d}{dt}\lambda_{qs2} + \omega_2\lambda_{ds2} \\
v_{ds2} &= R_{s2} \cdot i_{ds2} + \frac{d}{dt}\lambda_{ds2} - \omega_2\lambda_{qs2} \\
v_{qr} &= R_r \cdot i_{qr} + \frac{d}{dt}\lambda_{qr} + \omega_r\lambda_{dr} \\
v_{dr} &= R_r \cdot i_{dr} + \frac{d}{dt}\lambda_{dr} - \omega_r\lambda_{qr}
\end{aligned} \tag{5}
$$

where R, v, i and λ are the winding resistance, voltage, current and flow of control machine and power machine, respectively.

So, (5) can be written in matrix form as:

$$\dot{\Lambda} = \mathrm{V} - \mathrm{R} \cdot \mathrm{x} - \Omega \cdot \Lambda \tag{6}$$

where

$$
\Lambda = \begin{bmatrix} \lambda_{qs1} \\ \lambda_{ds1} \\ \lambda_{qs2} \\ \lambda_{ds2} \\ \lambda_{qr} \\ \lambda_{dr} \end{bmatrix}, \quad \mathrm{V} = \begin{bmatrix} \mathrm{v}_{qs1} \\ \mathrm{v}_{ds1} \\ \mathrm{v}_{qs2} \\ \mathrm{v}_{ds2} \\ \mathrm{v}_{qr} \\ \mathrm{v}_{dr} \end{bmatrix}, \quad \mathrm{x} = \begin{bmatrix} \mathrm{i}_{qs1} \\ \mathrm{i}_{ds1} \\ \mathrm{i}_{qs2} \\ \mathrm{i}_{ds2} \\ \mathrm{i}_{qr} \\ \mathrm{i}_{dr} \end{bmatrix} \tag{7}
$$

$$\mathrm{R} = \mathrm{diag}[\mathrm{R}_{s1} \quad \mathrm{R}_{s1} \quad \mathrm{R}_{s2} \quad \mathrm{R}_{s2} \quad \mathrm{R}_{r} \quad \mathrm{R}_{r}]$$

$$\Omega = \mathrm{diag}[\omega_1 \quad \omega_1 \quad \omega_2 \quad \omega_2 \quad \omega_r \quad \omega_r]$$

The stators and rotor flux vectors can be expressed as

$$\Lambda = \mathrm{L} \cdot \mathrm{x} \tag{8}$$

where

$$
\mathrm{L} = \begin{bmatrix} L_{s1} & 0 & 0 & 0 & L_m & 0 \\ 0 & L_{s1} & 0 & 0 & 0 & L_m \\ 0 & 0 & L_{s2} & 0 & -L_m & 0 \\ 0 & 0 & 0 & L_{s2} & 0 & L_m \\ -L_m & 0 & L_m & 0 & -L_r & 0 \\ 0 & L_m & 0 & L_m & 0 & L_r \end{bmatrix} \tag{9}
$$

Then by substituting (9) into (8) and rearrange the equations, (6) can be rewritten based on the derivations of currents. The Equivalent circuit of CDFIG is shown in Fig. 2.

If the stator of power machine is connected to grid, then the active and reactive powers transferred to grid are calculated as

$$\begin{aligned} \mathrm{P} &= \left(\frac{3}{2}\right)\left(\mathrm{v_{ds1}}\cdot\mathrm{i_{ds1}}+\mathrm{v_{qs1}}\cdot\mathrm{i_{qs1}}\right) \\ \mathrm{Q} &= \left(\frac{3}{2}\right)\left(\mathrm{v_{qs1}}\cdot\mathrm{i_{ds1}}-\mathrm{v_{ds1}}\cdot\mathrm{i_{qs1}}\right) \end{aligned} \tag{10}$$

The q-axis of the synchronous reference frame is aligned to the phase "a" of grid voltage so the d-component of the grid voltage is always zero. By considering the grid voltage reference frame, $v_{ds1}=0$. Then

$$\begin{aligned} P &= \left(\frac{3}{2}\right) v_{qs1}\cdot i_{qs1} \\ Q &= \left(\frac{3}{2}\right) v_{qs1}\cdot i_{ds1} \end{aligned} \tag{11}$$

In the second method, the MIMO system is considered solid and does not decoupled into SISO systems. Therefore, the multivariate system controller, known as the concentrated control method, must be programmed simultaneously.

Bristol proposed the idea of Relative Gain Array (RGA) to measure the interaction between input and output pairing variables [17]. The system transfer matrix must be multiplied by the element in the inverse of its transposed matrix to obtain RGA. If an element of the RGA matrix is close to the unit, it shows that the input and output variables are a suitable pair that can form a control loop. In other words, a small positive RGA element indicates that the dependence between the input and output variables is low [18].

$$\begin{aligned} &\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = G(s)\cdot\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \\ &G(s) = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \\ &RGA = G(s)\cdot\times\left(G(s)^T\right)^{-1} \end{aligned} \tag{12}$$

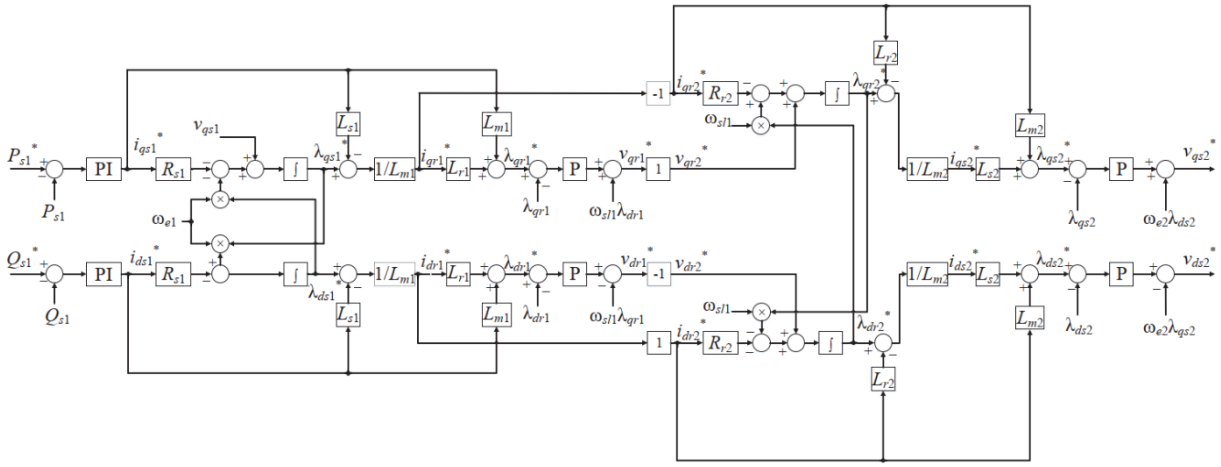


Fig. 3: Vector Control Method by Lipo [14].

Interaction between input and output

In multivariate systems, the interaction between input and output pairing variables is the main problem. This limitation can be overcome by planning suitable controllers for these systems in two main methods.

The first method is to try to convert the MIMO system to several SISO systems. Therefore, the system must be decoupled to eliminate interactions between input and output pairing variables.

By properly measuring the interaction between the input and output variables, the appropriate input and output pairing variables are selected. In this way, the system is decoupled into other SISO systems. A controller is then programmed for each loop, and finally the entire MIMO system is controlled.

In this paper the designing stages of multivariable control systems for non-minimum phase are studied.

The theories represented for decoupling multivariable systems can be applied only for minimum phase systems, and are not valid for non-minimum phase systems.

A. Before Grid Connection (Islanded mode)

Before connection of CDFIG output to grid or in islanded mode, the voltage magnitude, frequency and the voltage phase should be the same as grid ones.

When CDFIG is not connected to grid, the aim of controller is to control the terminal voltage of power machine. So, the inputs are voltage of control machine, v_{qs2} and v_{ds2}, and the outputs are voltage of power machine component, v_{qs1} and v_{ds1}.

$$\begin{bmatrix} v_{qs1} \\ v_{ds1} \end{bmatrix} = G_{IS}(s) \cdot \begin{bmatrix} v_{qs2} \\ v_{ds2} \end{bmatrix} \quad (13)$$

Then, using the system parameters listed in Table 1. RGA is calculated by (12) as follow:

$$RGA_{IS} = \begin{bmatrix} 0.9324 & 0.0676 \\ 0.0676 & 0.9324 \end{bmatrix}$$

The diagonal elements have the values close to unit and off-diagonal elements have the values close to zero in RGA_{IS}. Therefore, for the controller, the appropriate input and output pairs in (13) should be considered direct to direct voltage and the quadrature to quadrature voltage of the power and the control machine, respectively.

$$\begin{bmatrix} P_s \\ Q_s \end{bmatrix} = G_{GC}(s) \cdot \begin{bmatrix} v_{qs2} \\ v_{ds2} \end{bmatrix} \quad (14)$$

Then the RGA in this case is calculated by (12) as follow:

$$RGA_{GC} = \begin{bmatrix} 0.7487 & 0.2513 \\ 0.2513 & 0.7487 \end{bmatrix}$$

Therefore, since the RGA_{GC} is off-diagonal and not a near unitary matrix, the dependency between inputs and outputs is high.

Then the conventional cascaded PI control method isn't suitable for this machine.

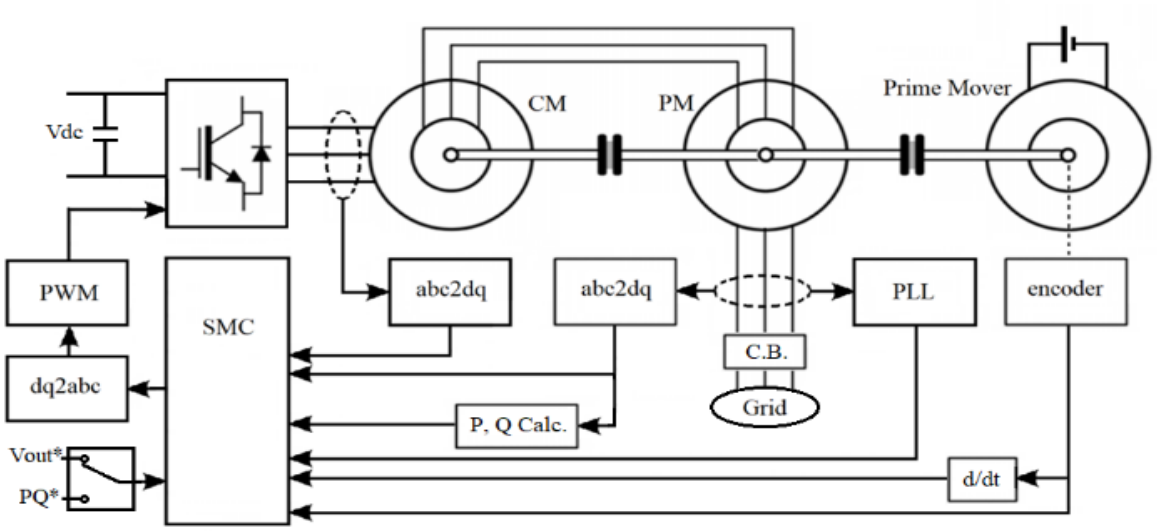


Fig. 4: Block diagram of SMC for CDFIG in islanded and grid connected mode.

So, in this case the cross correlation between inputs and outputs is low. In the other word, the v_{qs2} has low interaction on v_{ds1} and also like to v_{ds2} on v_{qs1}. Then the plant can be controlled by two decupled conventional PI controllers.

B. After Grid Connection

To connect the generator to grid, it is necessary to synchronize the power machine voltage with the grid, and then transferring the active and reactive with grid.

In [19], only deals with stand-alone mode for DFIG. However, the complexity of the equations and the dependence of the inputs and outputs are mostly determined in the grid-connected mode.

Calculation of RGA in grid connected mode confirms this.

For CDFIG system, when it is connected to grid, the input signals are terminal voltage of control machine, vqs2 and vds2, and the outputs are active and reactive power, P and Q.

Sliding Mode Control

In [14] a vector control for CDFIG is presented. As can be seen in Fig. 3, the controller is deeply dependent to parameters of machine. The controller is required the instantaneous rotor flux to be used as the feedback signal, which forces additional rotor-flux observer. Also, all of machine parameters are required in a recursive structure to generate the rotor reference currents and then the control machine reference currents using four P and two PI controllers can be calculated. So the tuning of PI coefficients is the main drawback of that method. The SMC is designed based on selecting a hyper plane in the state space or error space (called the sliding surface). If the state trajectory is confined to it, it will slide to the desired equilibrium point. By applying a control law based on sliding mode technique the system state is transferred to the sliding surface (in reaching mode) from an arbitrary initial condition and then, stays on (or close to) the sliding surface for all times and moves towards the equilibrium point (in sliding mode).

A. Before Grid Connection

The main target for SMC approach in islanded mode is to regulate the power machine voltage components in desired values, $v_{ds1} = v_{dref}$ and $v_{qs1} = v_{qref}$.

The sliding surfaces are selected in the integral forms as

$$S_d = e_d(\mathrm{t}) + K_d \cdot \int e_d\,(t)dt$$
$$S_q = e_q(\mathrm{t}) + K_q \cdot \int e_q\,(t)dt \quad (15)$$

where K_d and K_q are constant positive gains, and $e_d(\mathrm{t})$ and $e_q(\mathrm{t})$ are the differences between references and actual values of power machine voltage components, respectively.

$$e_d = v_{dref} - v_{ds1}$$
$$e_q = v_{qref} - v_{qs1}$$

Based on control design methods for sliding mode approach [19], the proper voltage references for VSI which feeds the control machine windings can be obtained as follows:

$$\begin{bmatrix}\mathrm{V_{qs2}}\\ \mathrm{V_{ds2}}\end{bmatrix} = -\mathrm{M}^{-1} \cdot \left(\mathrm{N} + \begin{bmatrix}\mathrm{K_d} \cdot \mathrm{sgn(S_d)}\\ \mathrm{K_q} \cdot \mathrm{sgn(S_q)}\end{bmatrix}\right) \quad (16)$$

where M and N can be calculated like as stated for grid connected mode in appendix.

The step by step design procedure will be explained for grid-connected mode and is ignored because of similarity.

B. After Grid Connection

After grid connection, the control targets are the active and reactive power transferred to grid via power machine. Also the control inputs are the components of control machine voltages, v_{qs2} and v_{ds2}.

The Time derivative of P and Q can be obtained as

$$\dot{P} = \left(\frac{3}{2}\right)\left(\dot{v_{qs1}} \cdot i_{qs1} + v_{qs1} \cdot \dot{\imath_{qs1}}\right)$$
$$\dot{Q} = \left(\frac{3}{2}\right)\left(\dot{v_{qs1}} \cdot i_{ds1} + v_{qs1} \cdot \dot{\imath_{ds1}}\right) \quad (17)$$

By considering $v_{qs1} = V_m$ and $\dot{v_{qs1}} = 0$, then

$$\dot{\mathrm{P}} = \left(\frac{3}{2}\right) V_m \cdot \dot{\imath_{qs1}}$$
$$\dot{Q} = \left(\frac{3}{2}\right) V_m \cdot \dot{\imath_{ds1}} \quad (18)$$

As the main task of the SMC approach, the active and reactive power must follow the reference values. Therefore, the sliding surface vector is given below:

$$\mathrm{S} = \begin{bmatrix}S_P\\ S_Q\end{bmatrix} \quad (19)$$

The sliding surfaces are assumed in the integral forms,

$$S_P = e_P(\mathrm{t}) + K_P \cdot \int e_P\,(t)dt$$
$$S_Q = e_Q(\mathrm{t}) + K_Q \cdot \int e_Q\,(t)dt \quad (20)$$

where K_P and K_Q are constant positive gains, and $e_P(\mathrm{t})$ and $e_Q(\mathrm{t})$ are the differences between references and actual values of active and reactive power respectively.

$$e_P = P_{ref} - P$$
$$e_Q = Q_{ref} - Q \quad (21)$$

when the system states achieve the desired surface, then we have:

$$S_P = \dot{S_P} = 0$$
$$S_Q = \dot{S_Q} = 0 \quad (22)$$

If the control laws are selected properly, we have

$$\dot{e_P} = -K_P \cdot e_P$$
$$\dot{e_Q} = -K_Q \cdot e_Q \quad (23)$$

It means that the errors will converge exponentially to zero.

According to (23)

$$\dot{S_P} = \dot{e_P} + K_P \cdot e_P = -\dot{P} + K_P \cdot (P_{ref} - P)$$
$$\dot{S_Q} = \dot{e_Q} + K_Q \cdot e_Q = -\dot{Q} + K_Q \cdot (Q_{ref} - Q) \quad (24)$$

Substituting (18) into (24) yields

$$\dot{\mathrm{S}} = \mathrm{E} + \mathrm{F} \cdot \mathrm{U} \quad (25)$$

With

$$\mathrm{E} = [E_{2*1}]$$
$$\mathrm{F} = [F_{2*2}]$$
$$\mathrm{U} = \begin{bmatrix}v_{qs2}\\ v_{ds2}\end{bmatrix} \quad (26)$$

where E and F are stated in appendix.

By applying Lyapunov theory in SMC method, the conditions of control law are derived and the state trajectory towards the desired behavior. Consider the quadratic function of Lyapunov as follows:

$$\mathrm{W} = \frac{1}{2}\mathrm{S^T S} \quad (27)$$

The time derivative of W of (22) is expressed as

$$\dot{W} = \frac{1}{2}\left(S^T\dot{S} + \dot{S}^T S\right) \quad (28)$$

The control law should be selected so that the time derivative of W is negative definite with $S \neq 0$. Thus, the control law is selected as follow:

$$\begin{bmatrix} V_{qs2} \\ V_{ds2} \end{bmatrix} = -F^{-1} \cdot \left(E + \begin{bmatrix} K_P \cdot \mathrm{sgn}(S_P) \\ K_Q \cdot \mathrm{sgn}(S_Q) \end{bmatrix}\right) \quad (29)$$

Where K_P and K_Q are positive control gains, $sgn(S_P)$ and $sgn(S_Q)$ are switching functions for active and reactive powers respectively. The block diagram of SMC for CDFIG is depicted in Fig. 4.

In order to decrease the chattering phenomena, the discontinues sign function can be replaced with saturation function. The reference voltage that is obtained from (29) can be used to drive the stator of control machine.

Simulation Results

Simulation of the proposed control strategy based on Fig. 4 is carried out by MATLAB/Simulink. Discrete model is used with a simulation time step of 20µs. The machine parameters are listed in Table 1.

Table 1: CDFIG parameters

Parameter	Value
V	220 V
Rs1	1.6 Ω
Rs2	1.6 Ω
Rr	3.2 Ω
p_1	1
p_2	1
L_{ls1}	0.004 H
L_{ls2}	0.004 H
L_{lr}	0.008 H
L_m	0.125 H

A. Before Grid Connection

The controller is used to control the stator voltage of power machine using the voltage source inverter which supplies the control machine windings. In this step, the first condition for grid connection, which is PM voltage equal to grid voltage, can be obtained.

The terminal voltage of the PM must be in the range of 220 volts and a frequency of 50 Hz and be in-phase with the grid voltage.

Terminal voltage of power machine in islanded mode with PI and SMC are shown in Fig. 5. Both of SMC and PI controller can control the voltage of CDFIG. But the dynamic response of PI controller is slower than SMC. As can be seen, the rise time of dynamic response with PI and SMC are 0.2 and 0.03 sec respectively.

B. After grid Connection

After connection of CFIG to grid, the controller must track the active and reactive power references. Considering the steady-state operation of the CDFIG with defined parameters as Table 1, the output active power reference is stepped up from 0.5 pu to 1.0 pu at t=1 s, and stepped down from 1.0 pu to 0.5 pu at t=1.5 s. As well as, the output reactive power reference is set to zero. Simulation results for active and reactive power control are shown in Figs. 5 and 6. For this condition, the obtained results are demonstrated by vector control and sliding mode control in Fig. 6 and Fig. 7, respectively. It can be seen, the output active and reactive powers follow the reference values in both methods, but SMC is so accurate and rapid. It is shown, the rise time of dynamic response with PI and SMC are 0.05 sec and 0.0003 sec respectively. As well as, the steady state error in two methods are zero. The overshoot in active power control with PI is about 10% and in SMC is also 10%. Furthermore, as can be seen in Fig. 6b and Fig. 7b, the fluctuation in reactive power when the active power reference is changed with step command is shown about 50% with PI controller while in SMC method no variation in reactive power is shown in the same condition of active power reference changes. It means the active and reactive power controllers are exactly decoupled in SMC method, while in vector control method the controllers are coupled, especially in transient regions.

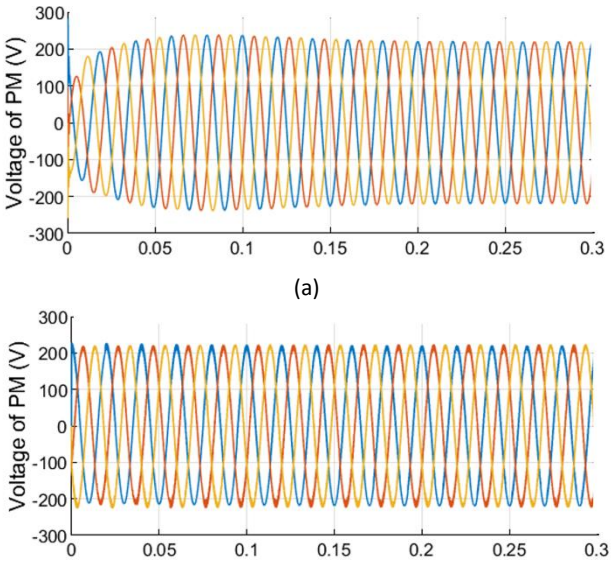


Fig. 5. Simulation results: Three phase Voltage of Power Machine Terminal. (a) Vector Control Method, (b) SMC Method.

C. Robustness against parameters uncertainties

In order to show the effectiveness of SMC against parameter uncertainties, a step change in rotor resistance is assumed and then the output behavior is depicted. Considering the active and reactive power is constant. At t=1.0 sec rotor resistance value is increased around 20 percent, from 3.2 Ω to 3.84 Ω. Simulation results are illustrated in Fig. 8. The output active and reactive power as well as their reference values, are depicted.

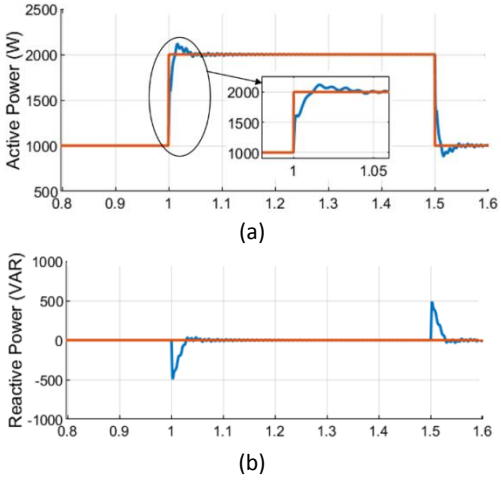


Fig. 6. Simulation results: Active and Reactive Power with step change in references values by Vector Control Method.

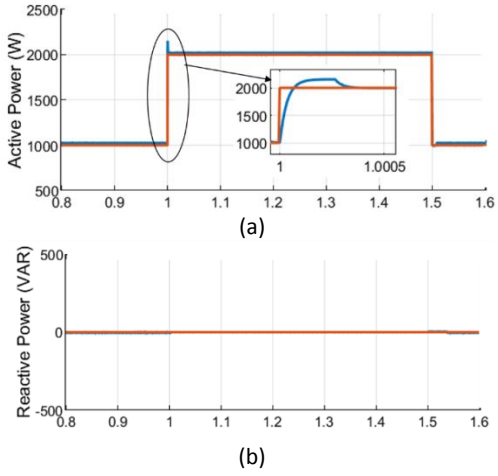


Fig. 7. Simulation results: Active and Reactive Power with step change in references values by SMC Method; a) Active power, b) Reactive power.

As can be seen, there very small fluctuations in the active and reactive power are shown, that verifies the performance of SMC.

Experimental Result

The performance of the proposed SMC is verified by a DSP-based prototype of CDFIG and its controller as depicted in Fig. 9. The practical setup consists of two DFIGs 370 W with cross interconnected rotor windings for implementation of CDFIG, one DC motor as a prime mover, voltage source inverter with its driver board to supplying the control machine, current and voltage sensor boards and a TMS320F28335 discrete signal processor board.

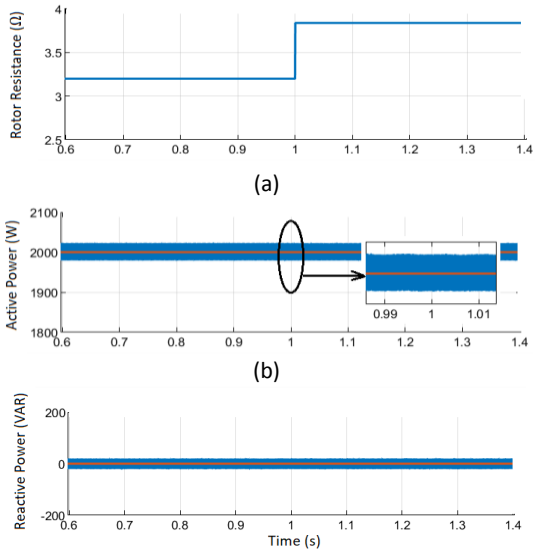


Fig. 8. Simulation results: P & Q control with increase in rotor resistance (SMC); a) Rotor resistance step change, b) Active power, c) Reactive power.

Two Hall-effect current sensors (LEM LA-55P) are measured the power machine currents, and a voltage sensor (LEM LV-25P) is calculated the line-to-line voltage. The analog second-order low pass filters are used to filtered the measured stator currents and voltage signals, with cut-off frequency of about 2.6 kHz, and converted to digital by 12-bit on-chip A/D converters.

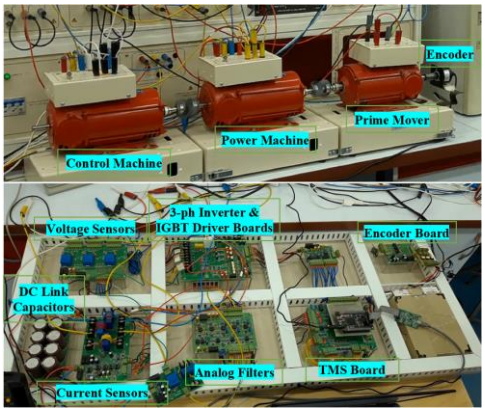


Fig. 9. Experimental setup of the grid connected CDFIG.

An incremental encoder with 1024 pulses per round connected to the DC motor shaft measures the rotor speed. The three phase inverter used in this setup includes six low loss IGBT switches kth123 (with 80 A, 1200 V ratings). Also, intelligent IGBT drivers, HCPL-316J, is used in this inverter which guarantee electrical separation between the power and control systems. The switching frequency of the inverter is selected as 10 kHz.

A. Before grid connection

The TMS board is used to control the stator voltage of power machine by a VSI which is connected to control machine. The terminal voltage of PM should be synchronized with the grid voltage. In order to ensuring the grid synchronization of power machine, Phase-Locked Loop (PLL) is used. Grid voltage is transformed to dq-synchronous rotating reference frame using Park transformation with estimated phase angle from the PLL output. Fig. 10 shows the grid and the CDFIG voltages before connection together.

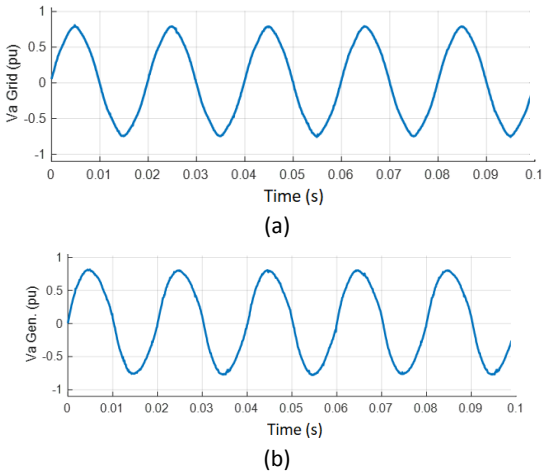


Fig. 10. Experimental results of Grid and Power machine terminal before connected together; (a) Grid Voltage, (b) Power Machine Voltage.

In order to show the effectiveness of SMC for output voltage control before connection to grid, the speed of prime mover is changed from 0.6 pu to 1 pu. Fig. 11 shows that, during the speed variation, the PM output voltage and frequency are controlled, but the Control machine current frequency decreases proportional to the CDFIG slip changes. Therefore, the robustness of proposed SMC against the rotor speed variation is verified.

B. Step change in output voltage in islanded mode

Controllability of terminal voltage of PM is verified by change in output reference voltage. Fig. 12 shows the output voltage of PM with a step change in references from -0.4 pu to 0.4 pu at t=10 s. In grid-tied mode, the active power reference is stepped up from 0.4 pu to 0.9 pu at t=2 s and stepped down from 0.9 pu to 0.4 pu at t=13 s. The reactive power reference is set to zero. Experimental results for vector control and sliding mode control are shown in Fig. 14.

C. Active and Reactive power control in grid connected mode

As can be seen, the rise time of dynamic response with PI and SMC are 2 and 0.2 sec respectively.

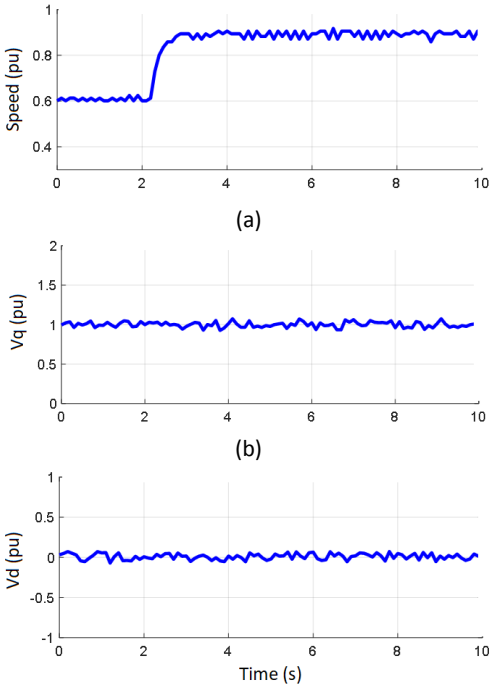


Fig. 11. Experimental result for speed change; (a) Speed of rotor, (b) Voltage of q-axis, (c) Voltage of d-axis.

This results show that the proposed SMC has very fast dynamic response in active power control. Also the interaction between active and reactive power is very low.

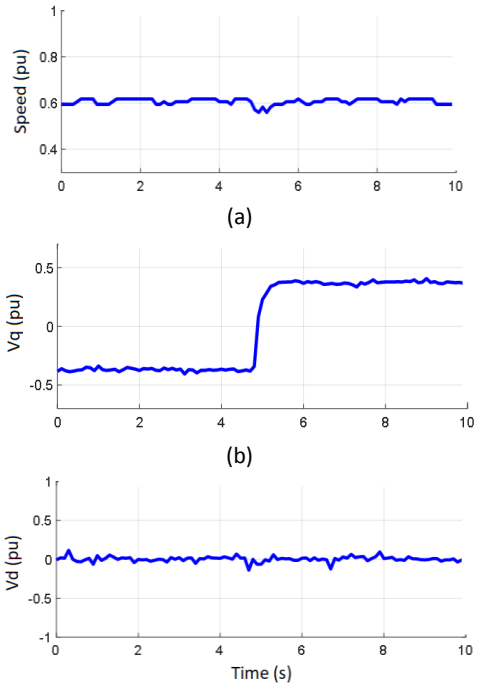


Fig. 12. Experimental result for voltage Ref. change; (a) Speed, (b) Voltage of q-axis, (c) Voltage of d-axis.

Fig. 13 shows the instantaneous output voltage of PM versus time.

The overall results are shown in the Table 2.

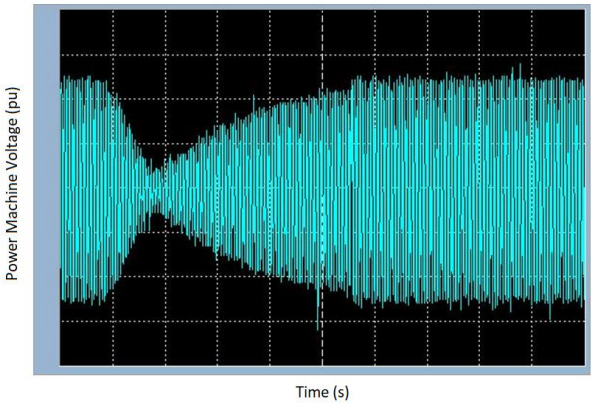


Fig. 13. Change in voltage reference from -0.4 pu to +0.4 pu.

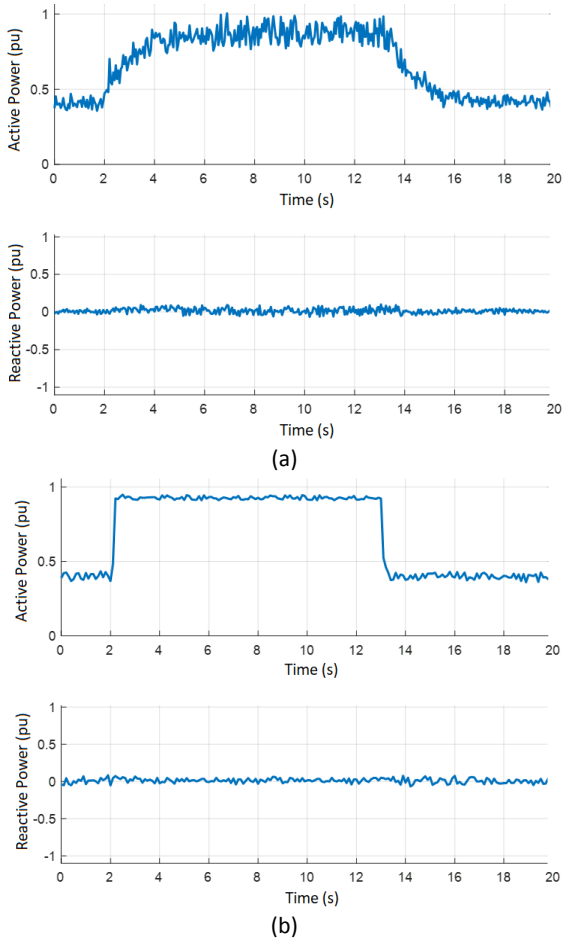


Fig. 14. Experimental results for changes active power reference at t=2s and t=13s; (a) vector control, (b) Sliding Mode Control.

Table 2: Simulation results

Parameter	PI	SMC
Voltage rise time (sim.)	0.2	0.03
Power rise time (sim.)	0.05	0.0003
Power rise time (exp.)	2	0.2

Conclusion

In this paper for the first effort, the RGA index is calculated to determine the interaction between the stator voltages of power machine and control machine in CDFIGs. The RGA index in islanded mode is near to diagonal matrix form and results a successful decoupled control for voltage control mode. While its results for grid-connected mode shows a huge interaction between all inputs and outputs, so the plant cannot be properly controlled by conventional PI controller. Therefore, the proposed sliding mode controller can be used to track the CDFIG targets in both islanded and grid-connected mode. As well as, the proposed method is robust to motor parameters uncertainties and has higher dynamic response in comparison with conventional vector control using PI regulators. Also, the effectiveness of the SMC method is verified by some simulation and experimental results, with variable wind speed and step changes in output targets.

Author Contributions

H. Zahedi, G.R. Arab Markadeh implemented the experimental setup and designed the simulations and experiments. S. Taghipoor, designed the CDFIG simulation model.

Acknowledgment

The authors wish to acknowledge the Shahrekord Univ. Supports.

Abbreviations

$v_{qs1}.v_{ds1}$	Voltages of d- and q- axis of power machine
$v_{qs2}.v_{qs2}$	Voltages of d- and q- axis of control machine
$v_{qr}.v_{dr}$	Voltages of d- and q- axis of rotor
$i_{qs1}.i_{ds1}$	Currents of d- and q- axis of power machine
$i_{qs2}.i_{qs2}$	Currents of d- and q- axis of control machine
$i_{qr}.i_{dr}$	Currents of d- and q- axis of rotor
$\lambda_{qs1}.\lambda_{ds1}$	Fluxes of d- and q- axis of power machine
$\lambda_{qs2}.\lambda_{ds2}$	Fluxes of d- and q- axis of control machine
$\lambda_{qr}.\lambda_{dr}$	Fluxes of d- and q- axis of rotor
$\omega_1.\omega_2$	Frequency of power and control machine
ω_m	Mechanical angular frequency
ω_r	Frequency of rotor flux

Appendix

A. *Current Model*

Current model for CDFIG can be described as

$$\dot{x} = A \cdot x + B \cdot u \qquad y = C \cdot x \qquad \text{(A1)}$$

where A, B and C are described in (A2), u is $[v_{qs2} \quad v_{ds2}]^T$ and $y = [P_s \quad Q_s]^T$ for grid connected and $y = [v_{qs1} \quad v_{ds1}]^T$ for isolated load.

$$\mathrm{A} = \frac{1}{L_a}\left[a_{ij}\right] \;:\; i.j = 1 \cdots 6$$

$$\mathrm{B} = \frac{1}{L_a}\begin{bmatrix} -L_m^2/Ls & 0 \\ 0 & L_m^2/Ls \\ L_b/Ls & 0 \\ 0 & L_b/Ls \\ L_m & 0 \\ 0 & -L_m \end{bmatrix} \tag{A2}$$

$$C = \left(2/3\right)v_{qs1}\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

By define $L_a = 2(L_r L_s - L_m^2)$, $L_b = 2L_r L_s - L_m^2$ and

$$\begin{gathered} a_{11} = a_{22} = a_{33} = a_{44} = -Rs * L_b/Ls \\ a_{12} = -a_{21} = -L_m^2 * wm - L_a * ws \\ a_{13} = -a_{24} = a_{31} = -a_{42} = L_m^2 * Rs/Ls \\ a_{14} = a_{23} = a_{32} = a_{41} = L_m^2 * wm \\ a_{34} = -a_{43} = -L_m^2 * wm - L_a * (2wm - ws) \\ a_{15} = a_{26} = -a_{35} = a_{46} = Lm * Rr \\ a_{16} = -a_{25} = a_{36} = a_{45} = -Lm * wm * L_a/Ls \\ a_{51} = a_{62} = -a_{53} = a_{64} = Lm * Rs \\ a_{52} = -a_{61} = -a_{63} = -a_{54} = Lm * Ls * wm \\ a_{55} = a_{66} = -Ls * Rr \\ a_{56} = -a_{65} = -L_a * (wm - ws) \end{gathered} \tag{A3}$$

The system output is the active and reactive power generated by CDFIG, that related to direct and quadrature current components of power machine. Therefore, the matrix C is shown in (A2).

B. Matrix E and F

By define $L_c = 4L_r L_s - L_m^2$

$$\mathrm{E} = \frac{1}{4L_s L_b}\begin{bmatrix} E_1 \\ E_2 \end{bmatrix}$$

$$\mathrm{F} = \begin{bmatrix} L_m^2 & 0 \\ 0 & L_m^2 \end{bmatrix} \tag{A3}$$

where

$$\begin{aligned} A_1 = {} & L_c L_s \omega_s i_{qs1} + L_c R_s i_{ds1} + L_m^2 L_s \omega_m i_{qs2} \\ & - L_m^2 R_s i_{ds2} + 2L_m L_b \omega_m i_{qr} \\ & + (L_m^2 L_s(\omega_m - \omega_s) + L_m L_s R_r) i_{dr} \end{aligned}$$

$$\begin{aligned} A_2 = {} & L_c L_s \omega_s i_{ds1} + L_c R_s i_{qs1} + L_m^2 L_s \omega_m i_{qs2} \\ & - L_m^2 R_s i_{ds2} + 2L_m L_b \omega_m i_{dr} \\ & + (L_m^2 L_s(\omega_m - \omega_s) + L_m L_s R_r) i_{qr} \end{aligned} \tag{A4}$$

References

[1] F. Blaabjerg, M. Liserre, K. Ma, “Power electronics converters for wind turbine systems,” IEEE Trans. Ind. Appl., 48(2): 708–719, 2012.

[2] B. Hopfensperger, D. J. Atkinson, R. A. Lakin, “Stator flux oriented control of a cascaded doubly-fed induction machine,” Proc. Inst. Elect. Eng.—Elect. Power Appl., 146(6): 597–605, 1999.

[3] T. D. Strous, H. Polinder, J. A. Ferreira, "Brushless doubly-fed induction machines for wind turbines: developments and research challenges," in IET Electric Power Applications, 11(6): 991-1000, 2017.

[4] E. Abdi, R. McMahon, P. Malliband, S. Shao, M. E. Mathekga, P. Tavner, S. Abdi, A. Oraee, T. Long, M. Tatlow, “Performance analysis and testing of a 250 kw medium-speed brushless doubly-fed induction generator,” IET Renew. Power Gener., 7(6): 631–638, 2013.

[5] M. Achkar, R. Mbayed, G. Salloum, S. Le Ballois, E. Monmasson, “Generic study of the power capability of a cascaded doubly fed induction machine,” International Journal of Electrical Power & Energy Systems, 86: 61-70, 2017.

[6] S. Ademi, M. G. Jovanović, M. Hasan, "Control of Brushless Doubly-Fed Reluctance Generators for Wind Energy Conversion Systems," in IEEE Transactions on Energy Conversion, 30(2): 596-604, 2015.

[7] Y. Tang, H. He, Z. Ni, J. Wen, X. Sui, “Reactive power control of grid-connected wind farm based on adaptive dynamic programming”, Neurocomputing, 125: 125–133, 2014.

[8] J. Chen, W. Zhang, B. Chen, Y. Ma, “Improved vector control of brushless doubly fed induction generator under unbalanced grid conditions for offshore wind power generation,” Energy Conversion, IEEE Transactions on, 31(1): 293-302, 2015.

[9] J.d.D.N. Ndongmo, G. Kenné, R.K. Fochie, A. Cheukem, H.B. Futsin, F.L. Lagarrigue, “A simplified nonlinear controller for transient stability enhancement of multimachine power systems using SSSC device”, Int. J. Electr. Power Energy Syst., 54: 650–657, 2014.

[10] G. Wang, R. Wai, Y. Liao, "Design of backstepping power control for grid-side converter of voltage source converter-based high-voltage dc wind power generation system," in IET Renewable Power Generation, 7(2): 118-133, 2013.

[11] H. E. Medouce, H. Benalla, A. Mehdi, A. Reama, "Sensorless direct power regulation by sliding mode approach of DFIG generator based wind energy system," in Proc. 2015 IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC), Rome: 1880-1885, 2015.

[12] Z. TIR, H. Rajeai, R. Abdessemed, “Analysis and vector control of a cascaded doubly fed induction generator in wind energy applications”. Revue des Energies Renouvelables: 347-358, 2010.

[13] M. E. Achkar, R. Mbayed, G. Salloum, N. Patin, S. Le Ballois, E. Monmasson, "Modeling and control of a stand alone cascaded doubly fed induction generator supplying an isolated load," 2015 17th European Conference on Power Electronics and Applications (EPE'15 ECCE-Europe), Geneva: 1-10, 2015.

[14] Z. S. Du, T. A. Lipo, "Dynamics and vector control of wound-rotor brushless doubly fed induction machines," in Proc. 2014 IEEE Energy Conversion Congress and Exposition (ECCE), Pittsburgh, PA: 1332-1339, 2014.

[15] J. Hu, J. Zhu, D. G. Dorrell, "A New Control Method of Cascaded Brushless Doubly Fed Induction Generators Using Direct Power Control," in IEEE Transactions on Energy Conversion, 29(3): 771-779, 2014.

[16] K. C. Wong, S. L. Ho, K. W. E. Cheng, “Direct voltage control for grid synchronization of doubly-fed induction generators,” Electr. Power Compon. Syst., 36(9): 960–976, 2008.

[17] M. Sadrnia, "A novel method for decoupling of non-minimum phase MIMO systems": 475-480, 2006.

[18] S. Z. Chen, N. C. Cheung, Y. Zhang, M. Zhang, X. M. Tang, "Improved Grid Synchronization Control of Doubly Fed Induction Generator Under Unbalanced Grid Voltage," in IEEE Transactions on Energy Conversion, 26(3): 799-810, 2011.

[19] S. Z. Chen, N. C. Cheung, K. C. Wong, J. Wu, "Grid Synchronization of Doubly-fed Induction Generator Using Integral Variable Structure Control," in IEEE Transactions on Energy Conversion, 24(4): 875-883, 2009.

Biographies

Hossein Zahedi Abdolhadi received the B.S. degree in Electrical Engineering from the University of Kashan, Kashan, Iran, in 2007 and the M.S. degree in Electrical Engineering from Tarbiat Modares University, Tehran, Iran, in 2010. He is currently working toward the PhD degree in the Faculty of Engineering, Shahrekord University, Shahrekord, Iran. His research interests include power converters, motor drivers and nonlinear control as well as control of power electronics and electric drives using microcontrollers and DSPs.

Gholamreza Arab Markadeh received the B.Sc., M.Sc., and Ph.D. degrees in Electrical Engineering from Isfahan University of Technology, Iran, in 1996, 1998, and 2005, respectively. He is currently an Associate Professor in the Faculty of Engineering, Shahrekord University. His fields of research include nonlinear control, power electronics, and variable-speed drives. He is the Editor-in-chief of Journal of Dam and Hydroelectric Powerplant. Dr. Arab Markadeh was the recipient of the IEEE Industrial Electronics Society IECON'04 best paper presentation award in 2004.

Samad Taghipour Boroujeni received the B.Sc, M.Sc., and Ph.D. degrees in electrical engineering from the Department of Electrical Engineering, Amirkabir University (Tehran Polytechnic), Tehran, Iran, in 2003, 2005, and 2009, respectively. In 2009, he joined the Department of Engineering, Shahrekord University, Shahrekord, Iran, as an Assistant Professor. Since 2016, he has been working as an Associate Professor. His research interests include modeling, design, analysis, optimization, and control of electrical machines, especially variable-speed generators.

How to cite this paper:
First Author, Second Author, and Third Author, "Real-time Implementation of Sliding Mode Control for Cascaded Doubly Fed Induction Generator in both Islanded and Grid Connected Modes," Journal of Electrical and Computer Engineering Innovations, 8(2): 285-296, 2020.

DOI: 10.22061/JECEI.2020.7361.384

URL: http://jecei.sru.ac.ir/article_1476.html

Journal of Electrical and Computer Engineering Innovations
(JECEI)
Journal homepage: http://www.jecei.sru.ac.ir

PAPER TYPE? (Research paper, short paper, Review paper *et al.*)

Instructions and Formatting Rules for Authors of Journal of Electrical and Computer Engineering Innovations, JECEI

F. Author, S. Author, T. Author

Affiliations of the Authors: (Department, Faculty, University(Institution), City, Country)

Article Info

Article History:
Received
Reviewed
Revised
Accepted

Keywords:
The author(s) shall provide up to 6 keywords to help identify the major topics of the paper

*Corresponding Author's Email Address:

Abstract

Background and Objectives: This section should be the shortest part of the abstract and should very briefly outline the following information: 1-What is already known about the subject, related to the paper in question. 2- What is not known about the subject and hence what the study intended to examine (or what the paper seeks to present). In most cases, the background can be framed in just 2–3 sentences, with each sentence describing a different aspect of the information referred to above; sometimes, even a single sentence may suffice. The purpose of the background, as the word itself indicates, is to provide the reader with a background to the study, and hence to smoothly lead into a description of the methods employed in the investigation.

Methods: The methods section is usually the second-longest section in the abstract. It should contain enough information to enable the reader to understand what was done, and how.

Results: The results section is the most important part of the abstract and nothing should compromise its range and quality. This is because readers who peruse an abstract do so to learn about the findings of the study. The results section should therefore be the longest part of the abstract and should contain as much detail about the findings as the journal word count permits.

Conclusion: This section should contain the most important take-home message of the study, expressed in a few precisely worded sentences. Usually, the finding highlighted here relates to the primary outcome measure; however, other important or unexpected findings should also be mentioned. It is also customary, but not essential, for the authors to express an opinion about the theoretical or practical implications of the findings, or the importance of their findings for the field. Thus, the conclusions may contain three elements: 1- The primary take-home message 2-The additional findings of importance 3-The perspective.

Introduction

This document provides an example of the desired layout for JECEI paper and can be used as a template for Microsoft Word versions 2003 and later. It contains information regarding desktop publishing format, type sizes, and typefaces. Style rules are provided to explain how to handle equations, units, figures, tables, abbreviations, and acronyms. Sections are also devoted to the preparation of appendixes, acknowledgments, references, and authors' biographies. For additional information including electronic file requirements for text and graphics, please refer to www.autjournal.com.

Doi:

Technical Work Preparation

Please use automatic hyphenation and check your spelling. Additionally, be sure your sentences are complete and that there is continuity within your paragraphs. Check the numbering of your graphics and make sure that all appropriate references are included.

A. Template

This document may be used as a template for preparing your technical work.

B. Format

If you choose not to use this document as a template, prepare your technical work in single-spaced, double-column format, on paper 21.6×27.9 centimeters (8.5×11 inches or 51×66 picas). Set top and bottom margins to 25 millimeters (0.98 inch) and left and right margins to about 20 millimeters (0.79 inch). Do not violate margins (i.e., text, tables, figures, and equations may not extend into the margins). The column width is 82 millimeters (3.2 inches). The space between the two columns is 6 millimeters (0.24 inch). Paragraph indentation is 4.2 millimeters (0.17 inch). Use full justification. Use either one or two spaces between sections, and between text and tables or figures, to adjust the column length.

C. Typefaces and Sizes

Please use a proportional serif typeface such as Calibri and embed all fonts. Table 1 provides samples of the appropriate type sizes and styles to use.

D. Section Headings

A primary section heading is enumerated by a Roman numeral followed by a period and is centered above the text. A primary heading should be in capital letters.

A secondary section heading is enumerated by a capital letter followed by a period and is flush left above the section. The first letter of each important word is capitalized and the heading is italicized.

A tertiary section heading is enumerated by an Arabic numeral followed by a parenthesis. It is indented and is followed by a colon. The first letter of each important word is capitalized and the heading is italicized.

A quaternary section heading is rarely necessary, but is perfectly acceptable if required. It is enumerated by a lowercase letter followed by a parenthesis. It is indented and is followed by a colon. Only the first letter of the heading is capitalized and the heading is italicized.

E. Figures and Tables

Figure axis labels are often a source of confusion. Try to use words rather than symbols. As an example, write the quantity "Magnetization," or "Magnetization, *M*," not just "*M*." Put units in parentheses. Do not label axes only with units. As in Fig. 1, write "Magnetization (kA/m)" or "Magnetization (kA·m^{-1})," not just "kA/m." Do not label axes with a ratio of quantities and units. For example, write "Temperature (K)," not "Temperature/K." Figure labels should be legible, approximately 8- to 10-point type.

Large figures and tables may span both columns, but may not extend into the page margins. Arrange these one column figures and tables at either top or end of a page, or at the end of the paper right before the references. Figure captions should be below the figures; table captions should be above the tables. Do not put captions in "text boxes" linked to the figures. Do not put borders around your figures. Use Insert | Reference | Caption to number your tables and figures, and use Insert | Reference | Cross- reference to refer to their numbers.

Table 1: Samples of Calibri sizes and styles used for formatting a pes technical work

Point Size	Purpose in Paper	Special Appearance
9	Table text, figure text footnotes, subscripts, superscripts, references, bio, Figure caption, keywords	Table Title
10	Body text, equations, author affiliation, abstract	*Subheadings*
11		**Section Titles**
12	***Author Name***	

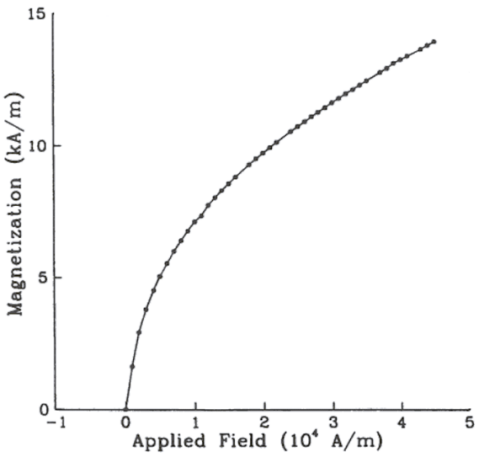


Fig. 1: Magnetization as a function of applied field. (Note that there is a colon after the figure number followed by two spaces.)

All figures and tables must appear near, but not before, their first mention in the text. Use the abbreviation "Fig. 1," even at the beginning of a sentence.

To insert images in Word, use Insert | Picture | From File.

F. Numbering

Number reference citations consecutively in square

brackets [1]. The sentence punctuation follows the brackets [2]. Multiple references [2], [3] are each numbered with separate brackets [1][1]-[2]. Refer simply to the reference number, as in [2]. Do not use "Ref. [2]" or "reference [2]" except at the beginning of a sentence: "Reference [2] shows…."

Number footnotes separately with superscripts (Insert | Footnote). Place the actual footnote at the bottom of the column in which it is cited. Do not put footnotes in the reference list. Use letters for table footnotes.

Use Arabic numerals for figures and Roman numerals for tables. Appendix figures and tables should be numbered consecutively with the figures and tables appearing in the rest of the paper. They should not have their own numbering system.

G. Units

Metric units are preferred for use in IEEE publications in light of their global readership and the inherent convenience of these units in many fields. In particular, the use of the International System of Units is advocated. This system includes a subsystem of units based on the meter, kilogram, second, and ampere (MKSA). British units may be used as secondary units (in parentheses). An exception is when British units are used as identifiers in trade, such as 3.5-inch disk drive.

H. Math and Equations

Number equations consecutively with equation numbers in parentheses flush with the right margin, as in (1). First use the equation editor to create the equation. Then select the "Equation" markup style. Write the equation number in parentheses using Insert | Caption.

Use the Microsoft Equation Editor for all math objects in your paper (Insert | Object | Create New | Microsoft Equation *or* MathType Equation). "Float over text" should *not* be selected.

To make your equations more compact, you may use the slash (/), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Use parentheses to avoid ambiguities in denominators. Number equations consecutively with equation numbers in parentheses flush with the right margin, as in (1). Be sure that the symbols in your equation have been defined before the equation appears or immediately following. Italicize symbols (*T* might refer to temperature, but T is the unit Tesla).

Use Insert | Reference | Caption to number equations. Refer to "(1)," not "Eq. 1" or "equation (1)," except at the beginning of a sentence: "Equation (1) is ...". Punctuate equations when they are part of a sentence, as in

$$\int_0^{r_2} F(r,\varphi)\, dr\, d\varphi = [\sigma r_2 / (2\mu_0)] \cdot \int_0^{\infty} \exp(-\lambda\,|\,z_j - z_i\,|)\,\lambda^{-1}\, J_1(\lambda r_2)\, J_0(\lambda r_i)\, d\lambda \quad (1)$$

Use two column tables to locate equations and their numbers properly in one line, as follows:

$$I_F = I_B = -I_C = A^2 I_{A1} + A I_{A2} + I_{A0} = \frac{-J\sqrt{3}E_A}{Z_1 + Z_2} \quad (2)$$

where I_F is the fault current. Be sure that the border is off.

Results and Discussion

The Results section should briefly present the experimental data in text, tables or figures. Tables and figures should not be described extensively in the text.

The Discussion should focus on the interpretation and the significance of the findings with concise objective comments that describe their relation to other work in the area. It should not repeat information in the results. The final paragraph should highlight the main conclusion(s), and provide some indication of the direction future research should take.

Conclusion

As the Conclusion section is the most important element of a manuscript, so it must be more expanded scientifically and contently at least half a page length.

Example:

In this study, a forecast model was developed to determine the generation of MSW in the municipalities of the CCS, Chiapas State, Mexico. A MLR was used to obtain the forecast model with social and demographic explanatory variables. Two forecast models were presented and analyzed, with variables that met the multicollinearity test. The most important variables to predict the rate of MSW generation in the study area were the population of each municipality (XPop), the population born in another municipality (XPbam) and the population density (XPd). XPop is the most influential explanatory variable of waste generation, particularly it is related in a positive way. XPbam is less related to waste generation. XPd is the variable that least influences waste generation prediction; in addition, it can present problems of correlation with other explanatory variables. Although other variables, such as daily per capita income (XDpi) and average schooling (XAs), are very important, they do not seem to have an effect on the response variable in this study. The user of this forecast model should use model 2, since it is the one with the highest parsimony (it uses fewer variables); R^2_{adj}, MAPE, MAD and RMSE values indicated high influence on the explained phenomenon and high forecasting capacity. Additionally, it is important to mention that when using the models proposed for forecasting purposes, it is necessary to make a

transformation in the explanatory and response variables (use inverse of natural logarithm). The inferences made on the municipalities of the study area showed that, except in some municipalities, the MSW generation rate usually presented a gradual increase with respect to population growth and with respect to the number of inhabitants that were born in another entity (migration). Finally, this study can be a solid basis for comparison for future research in the area of study. It is possible to use different mathematical models such as artificial neural network, principal component analysis, time-series analysis, etc., and compare the response variable or the predictors.

Author Contributions

Each author role in the research participation must be mentioned clearly.

Example:

A. Mahboobi, B. Bagheri, and C. Ahmdi designed the experiments. A. Mahboobi collected the data. A. Mahboobi carried out the data analysis. A. Mahboobi, B. Bagheri, and C. Ahmdi interpreted the results and wrote the manuscript.

Acknowledgment

The following is an example of an acknowledgment. (Please note that financial support should be acknowledged in the unnumbered footnote on the title page.)

The author gratefully acknowledges the IEEE I. X. Austan, A. H. Burgmeyer, C. J. Essel, and S. H. Gold for their work on the original version of this document.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Abbreviations

Define less common abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, ac, dc, and rms do not have to be defined. Do not use abbreviations in the title unless they are unavoidable.

Example:

MS	Multispectral
SMF	Spectral Matched Filter
SAM	Spectral Angle Mapper
MSD	Matched Subspace Detector
OSP	Orthogonal Subspace Projection
CEM	Constrained Energy Minimization
ASD	Adaptive Subspace Detector
STD	Sparsity Based Target Detector
KSAM	Kernel Based SAM
DTD	Difference Based Target Detection
AP-CR	Attribute Profile Based Collaborative Representation
ROC	Receiver Operating Characteristic
MS	Multispectral
SMF	Spectral Matched Filter
SAM	Spectral Angle Mapper
MSD	Matched Subspace Detector
OSP	Orthogonal Subspace Projection
CEM	Constrained Energy Minimization
ASD	Adaptive Subspace Detector
STD	Sparsity Based Target Detector
KSAM	Kernel Based SAM
DTD	Difference Based Target Detection
AP-CR	Attribute Profile Based Collaborative Representation
ROC	Receiver Operating Characteristic
MS	Multispectral
SMF	Spectral Matched Filter
SAM	Spectral Angle Mapper
MSD	Matched Subspace Detector
OSP	Orthogonal Subspace Projection
CEM	Constrained Energy Minimization
ASD	Adaptive Subspace Detector
STD	Sparsity Based Target Detector
KSAM	Kernel Based SAM

References

References are important to the reader; therefore, each citation must be complete and correct. There is no editorial check on references; therefore, an incomplete or wrong reference will be published unless caught by a reviewer or discusser and will detract from the authority and value of the paper. References should be readily available publications. List only one reference per reference number. If a reference is available from two sources, each should be listed as a separate reference. Give all authors' names; do not use *et al.*

Samples of the correct formats for various types of references are given below.

Periodicals:

[1] J. F. Fuller, E. F. Fuchs, K. J. Roesler, "Influence of harmonics on power distribution system protection," IEEE Trans. Power Deliv., 3(2): 549-557, 1988.

Books:

[2] E. Clarke, Circuit Analysis of AC Power Systems, vol. I. New York: Wiley: 81, 1950.

Technical Reports:

[3] E. E. Reber, R. L. Mitchell, C. J. Carter, "Oxygen absorption in the Earth's atmosphere," Aerospace Corp., Los Angeles, CA, Tech. Rep. TR-0200 (4230-46)-3, Nov. 1968.

[4] S. L. Talleen. (1996, Apr.). The Intranet Architecture: Managing information in the new paradigm. Amdahl Corp., Sunnyvale, CA.

Papers Presented at Conferences (Unpublished):

[5] D. Ebehard, E. Voges, "Digital single sideband detection for interferometric sensors," presented at the 2nd Int. Conf. Optical Fiber Sensors, Stuttgart, Germany, 1984.

[6] Process Corp., Framingham, MA. Intranets: Internet technologies deployed behind the firewall for corporate productivity. Presented at INET96 Annu. Meeting.

Papers from Conference Proceedings (Published):

[7] J. L. Alqueres, J. C. Praca, "The Brazilian power system and the challenge of the Amazon transmission," in Proc. IEEE Power Engineering Society Transmission and Distribution Conf.: 315-320, 1991.

Dissertations:

[8] S. Hwang, "Frequency domain system identification of helicopter rotor dynamics incorporating models with time periodic coefficients," Ph.D. dissertation, Dept. Aerosp. Eng., Univ. Maryland, College Park, 1997.

Standards:

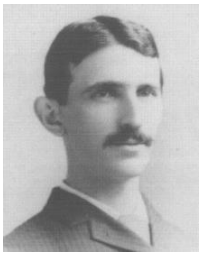
[9] IEEE Guide for Application of Power Apparatus Bushings, IEEE Standard C57.19.100-1995, Aug. 1995.

Patents:

[10] G. Brandli and M. Dick, "Alternating current fed power supply," U.S. Patent 4 084 217, Nov. 4, 1978.

Biographies

A technical biography for each author may be included, but without any title, as it is seen herein. It should begin with the author's name (as it appears in the byline). A photograph and an electronic file of the photo should also be included for each author. The photo should be black and white, glossy, and 3.0 centimeters (1.18 inches) wide by 3.8 centimeters (1.5 inches) high. The head and shoulders should be centered, and the photo should be flush with the left margin. The following is an example of the text of a technical biography:



Nikola Tesla (M'1888, F'17) was born in Smiljan in the Austro-Hungarian Empire, on July 9, 1856. He graduated from the Austrian Polytechnic School, Graz, and studied at the University of Prague. His employment experience included the American Telephone Company, Budapest, the Edison Machine Works, Westinghouse Electric Company, and Nikola Tesla Laboratories. His special fields of interest included high frequency. Tesla received honorary degrees from institutions of higher learning including Columbia University, Yale University, University of Belgrade, and the University of Zagreb. He received the Elliott Cresson Medal of the Franklin Institute and the Edison Medal of the IEEE. In 1956, the term "tesla" (T) was adopted as the unit of magnetic flux density in the MKSA system. In 1975, the Power Engineering Society established the Nikola Tesla Award in his honor. Tesla died on January 7, 1943.

How to cite this paper:

F. Author, S. Author, T. Author, "Instructions and formatting rules for authors of journal of electrical and computer engineering innovations, JECEI," J. Electr. Comput. Eng. Innovations, ×(×): ×××-×××, ××××.

DOI:

URL: http://jecei.srttu.edu/journal/authors.note