Quarterly Journal of

Optimization In Soft Computing

Vol. 1, No. 1, Autumn 2023

Journal of Optimization of Soft Computing (JOSC)

Vol. 1, Issue 1, pp: (1-11), September-2023

Journal homepage: https://sanad.iau.ir/journal/josc

Research paper

A New Identity-Based Encryption Scheme Using Blockchain for Electronic Health System

Fatemeh Alidadi Shamsabadi[1], Shaghayegh Bakhtiari Chehelcheshmeh[2*] and Majid Alipour[3]

[1,2,3] *Department of Computer, Faculty of Engineering, Shahrekord Branch, Islamic Azad University, Shahrekord, Iran.*

Article Info

Article History:
Received :2023/8/12
Revised :2023/8/21
Accepted 2023/9/1

Keywords:
Electronic Health System, Identity-Based Encryption, Blockchain, Cloud Computing.

**Corresponding Author's Email Address: sh.bakhtiari@iaushk.ac.ir*

Abstract

With the development of information technology, electronic health (e-health) systems are used as a common approach to recording patients' medical data. Given that medical information is an essential asset of individuals and the health system, there are severe concerns about secure sharing and preserving this information's privacy. In this paper, we propose a new identity-based encryption (IBE) method by bilinear pairings to ensure the confidentiality of patients' data and their privacy in cloud health systems; in this way, this plan also provides authentication for users using challenge-response mechanisms. In addition, the proposed scheme, using blockchain techniques, ensures integrity and precise access control for shared data. The correctness of the proposed protocol is verified, and its security is formally proven in the standard model. The implementation of our scheme is performed in Java, and the results show that the proposed scheme reduces the computational overhead compared to previous similar methods.

1. Introduction

Today, e-health systems are rapidly evolving and becoming operational because these systems increase the quality of health care by enabling the exchange of information between health centers and joint treatment decisions. They also prevent unnecessary tests, misdiagnosis, and repetitive treatments by strengthening links between healthcare institutions [1, 2].

One of the main challenges in developing and advancing electronic health systems is providing confidentiality service and privacy. Since any attack on these systems will lead to irreparable damage, in recent years, many schemes have been proposed to ensure the security and confidentiality of electronic health systems.

Many of these methods are based on public-key protocols. In traditional systems, the public key infrastructure (PKI) requires a trusted third party to issue digital certificates so that users can verify the authenticity of other users' public keys. In addition to increasing the system's complexity, this increases the cost of electronic health systems [3]. In an identity-based encryption system, the identity of users (such as email addresses) is used to generate the keys of users. Thus, there is no need to issue and verify a digital certificate in an identity-based cryptography system [4]. The use of identity-based encryption in electronic health systems makes users comfortable. And also significantly reduces the costs of these types of systems.

Re-encryption proxy is another encryption method that converts ciphertext using the first user's identity to another ciphertext using the second user's identity. Therefore, in addition to preserving the confidentiality of the message, the second user can decrypt the ciphertext. Since the electronic health system needs to provide encrypted information by the patient to other authorized departments using the ID of the medical team, the use of re-encryption proxies in these systems facilitates the patient service process. Also, it reduces the cost of cryptography [3].

In addition to security issues, sharing and synchronizing sensitive information in e-health systems is also very important. Patient information held by the medical team is not readily available to other health centers. As a result, duplicate medical records are created. In addition to being costly for the health care system, this poses other processes for patients, which wastes their time and unnecessarily extra costs. Blockchain, meanwhile, allows healthcare professionals to share this information through a chain of non-tampering blocks rather than storing patient records in separate databases [5, 6].

However, using blocks as a storage place for all information requires a lot of storage space and is expensive. Therefore, the content of patients' data can be stored in encrypted form in the cloud medical providers. Only the access address and a brief description of that information in the cloud providers can be recorded in the blocks to be accessed when needed [7].

Therefore in this research, a new identity-based encryption scheme is presented; Solves issues related to digital certificate management and public key infrastructure complexity. The proposed IBE scheme uses bilinear pairings only in the decryption phase, and the pairing parameter used in the encryption phase is pre-calculated; for this reason, using the proposed IBE reduces the computational overhead in the electronic health system.

The proposed design also can be used in re-encryption and blockchain technology. Due to the high volume of information, the content of patients' medical records is stored as encrypted in cloud storage systems. The access address of each of them is placed in a blockchain to share better and manage this data.

In addition, the effective combination of this scheme and re-encryption proxy provides confidentiality, reduces costs, and increases accessibility in the health system. In this paper, the proposed scheme is proved both in the random oracle and standard models. Also, the implementation results show that the proposed method reduces the running time compared to similar schemes.

The subsequent sections of the paper are organized as follows: In Section 2, an overview of previous schemes is provided. Section 3 describes the requirements for the proposal. In Section 4, the system is presented. In Section 5, the security and efficiency of the scheme are proved and evaluated. Finally, Section 6 concludes the paper.

2. Related Works

E-health is an electronic and communication process aimed at supporting patients' health care, introduced in 1999 [8]. The Electronic Medical Record (EMR) is kept confidential in these systems. E-health records contain information about the health of individuals during their lifetime, which serves as a source of information for physicians, health care providers, and research centers. Using electronic health systems, physicians can easily change health records. Patients also have easy access to their medical records through user interfaces designed to view medical information; In addition, healthcare providers can access patient information in critical situations [9, 3].

However, confidentiality and privacy are the main challenges in developing the health system to record patients' health records [10].

There are currently many suggestions for using cryptography to provide confidentiality in health systems, including symmetric and Asymmetric key methods similar to the anonymous ID technique. A common belief for creating confidentiality and privacy in e-health systems is to encrypt e-health documents in these systems [3].

Thus, data, identifiers, and attribute data keys (metadata) must be encrypted before being stored in an authentication center or resource in the cloud. In addition, establishing access control and key management issues and the cost of its implementation are always the most vital issues of health systems. Hence, cryptographic techniques can implement secure access control mechanisms or key management properties [11–13].

Despite the capabilities of IBE in recent years, very few articles have used these techniques to create confidentiality and privacy in electronic cloud health systems. Therefore, in this section, works close to the subject are expressed.

Benaloh et al. (2009) explored the challenges of protecting patients' privacy in the electronic health record system, arguing that security in such systems should be provided using encryption and access control. They also discussed using cryptography to record electronic medical records, ensuring confidentiality by a patient-controlled cryptography model [14]. In this scheme, the patient can generate and store the cryptographic key; So that the patient's privacy in the computer data center is not compromised. In the same year, Xue et al. proposed an authentication system using identity-based encryption to protect the confidentiality of electronic health networks [15].

In 2011, Barua et al. proposed a plan for the confidentiality and control of patient-centered access to electronic health systems in the cloud [16]. In this scheme, to ensure the privacy of personal health information, the EDPAC technique has been introduced, which enables the system to give data requesters different levels of access based on their role in the system.

In 2012, Guo et al. introduced an authentication system with feature-based confidentiality in the e-health system [17]. In this scheme, to protect patients' personal information while receiving medical services, a framework called PAAS has been proposed to authenticate the attribute of users in the process of authentication of users in the electronic health system to preserve privacy.

In 2013, Li et al. proposed the scalable and secure sharing of personal health records in the cloud using feature-based encryption [18]. This design presents a new patient-centric model with a set of data access control mechanisms for storing personal health documents using the feature-based encryption technique in semi-trusted providers.

In 2016, Yan et al. proposed a new access plan to implement secure access control of personal health documents in electronic health systems based on the original feature-based encryption scheme [19].

According to studies, the schemes [3, 7] increase access to medical records and are considered to be the most relevant plans to the proposed health system. Now, these schemes will be reviewed.

In 2017, Wang et al. presented a new identity-based encryption scheme for utilization in the health system. They explained how to use this scheme to reduce the disadvantages of the methods used in the e-health system. This identity-based encryption scheme can be used in a re-encryption proxy. However, the extensive use of bilinear pairing functions and extended general parameters increases the computational cost in the encryption and decoding phases [3].

In 2018, Wang and Song implemented a hybrid encryption scheme in blockchain technology to share e-health records securely. The use of two attribute-based signature schemes and identity-based encryption in this method dramatically increases the computational overhead and overshadows the scalability of this system [7].

3. Prerequisites for the Proposed Scheme

In this paper, we first introduce a new method for identity-based encryption. The proposed method is then combined with a re-encryption proxy, blockchain technology, and cloud storage systems for electronic health systems. For this purpose, cryptographic methods based on identity are examined in this section, and other concepts related to the proposed design are described.

3.1. Identity-Based Encryption

As the name implies, identity-based encryption uses the user ID (such as Email Address) as the public key in this type of encryption. In this encryption method, Key Generation Center (KGC) is responsible for the authentication of users. The identifiers are registered; KGC creates the private key corresponding to the ID and delivers it to the users. It also creates some public parameters and makes them known to everyone.

The user who intends to send the message first encrypts the message with the recipient's public key obtained directly from her ID and the public parameters provided by the key generation center. The message recipient decrypts the ciphertext using its private key and the public parameters from the key generation center [20]. Shamir first proposed this cryptographic design in 1984 as an identity-based signature (IBS) method [21].

However, Shamir failed to introduce identity-based encryption until Boneh and Franklin 2001 introduced the first practical structure of identity-based encryption based on bilinear pairing groups with proof in the random oracle model [4]. Due to the reasonable length of the key and the cost of computations for the key generation center, this design became the basis for creating IBE methods.

In 2003, Boneh and Franklin introduced another identity-based encryption. This design was presented with slight changes from their first design, such as asymmetric bilinear pairing and the use of two hash functions as a public parameter of the key generation center with proof in the random oracle model [22].

Subsequently, other methods with relatively different structures were proposed, among which the IBE and IBS methods of Sakai and Kasahara 2003 [23] and the IBE method of Boneh and Boyen 2004 [24] were more effective. The reason for this better efficiency was not to use bilinear pairing at encryption, which reduced costs.

In 2006, Gentry introduced an identity-based encryption scheme based on the Boneh and Boyen method, which required the calculation of further pairings in the form of parameters pre-calculated by the key generation center [25]. Gentry design has more resistance than Boneh and Boyen's scheme. Still, it is less efficient than the Sakai and Kasahara designs and the non-random oracle version of Boneh and Boyen, and it also requires complex assumptions [20].

In 2011, Boneh and Boyen introduced an identity-based encryption system. Their system was not based on the random oracle model, and it provided resistance against the chosen-identity attack. This structure did not need to calculate bilinear pairing in the encryption phase. It is only required to calculate two pairing functions in the decryption phase [26]. In 2010, Galindo proposed an identity-based encryption design. It provided chosen-ciphertext security with constant-size ciphertexts under computational Diffie-Hellman's bilinear assumptions in the standard model [27].

In 2015, Park et al. proposed an identity-based encryption scheme with chosen-ciphertext security. They aimed to obtain tight security in the bilinear Diffie Hellman problem [28]. In 2015, Susilo et al. presented an identity-based encryption scheme with a dynamic threshold and constant-size ciphertext. In their design, the sender, by selecting the number of receivers, provides constant-size ciphertext so that it is possible to decrypt only in the presence of all receivers [29].

In 2017, Bakhtiari and Hosseinzadeh introduced an identity-based encryption method that, in addition to encryption, can also use for signing messages. They used their method for proposing a certificateless authentication scheme [30]. In the same year, Wang et al. introduced the e-health system using an identity-based cryptographic scheme that was efficient [3]. This method is optimized for use in re-encryption proxies. The disadvantages of this method are the long public parameters and the use of many bilinear pairing functions in encryption and decryption.

3.2. Re-Encryption Proxy

The concept of re-encryption proxy was introduced in 1998 by Blaze et al. [31]. In this scheme, a semitrusted proxy with information such as a re-encryption key converts the ciphertext into another ciphertext using the public key of the first delegate; So that it is possible to decryption the ciphertext by the private key of the second delegate. In this scheme, the proxy will not access the message's plaintext.

3.3. Blockchain Technology

The first blockchain was introduced in 2008 by Nakamoto to create a consistent database among all members of a decentralized network [32]. Here is an example of its application in the health system to better understand blockchain technology.

People often go to different hospitals and medical centers during their lives, and each time they go, they provide specific and various medical information to these institutions. Hospitals must share patients' medical information confidentially with other authorized institutions to provide accurate, fast, and effective medical services [33]. In addition, they need to address concerns about patients' privacy, mistrust between healthcare providers, scalability, and how to control access to medical information [34, 35].

For this purpose, blockchain technology is used as a suitable solution. Blockchain technology allows the information stored in the blocks to be shared among all network members, making it almost impossible to manipulate the recorded data using cryptographic methods. In addition, access control can be achieved more efficiently by implementing blockchain. When this system is used for large volumes of data, health professionals will be well aware of how accurate the management of this new trend can be [36, 5].

4. Proposed Scheme

In this section, a new identity-based encryption scheme based on bilinear pairings is first proposed. Then, the model of the re-encryption proxy system is presented according to the proposed cryptographic scheme. Finally, the cloud health system for medical information sharing using blockchain technology is described. Table 1 describes the symbols used in the proposed scheme.

Table 1: Symbols of the proposed scheme

Symbol	Definition
$h()$	Hash function
G, G_T	Cyclic group
$e: G \times G \to G_T$	Function bilinear pairing on the elliptic curve
ID_i	Identity-related to the ith user
$msk = \{\}$	KGC's private key
g_1, g_2	Public parameters
Z_P	The prime group to measure p
Z^*_p	The prime group, which contains all positive integers, is smaller than p.
g	Generator of group Z^*_p
d_i	Private key related to the ith user
m	Plaintext
CT	Ciphertext
SE	Symmetric cryptographic element
δ	One-time digital signature pattern

4.1. Proposed Identity-Based Encryption

The focus of the proposed identity-based encryption scheme is to maintain flexibility while providing security so that the proposed IBE is independent in other applications. In the proposed IBE, the computational overhead is significantly reduced by shortening the parameters used in the cryptographic phases and pre-calculated pairing in the encryption phase. In addition, the strength of the proposed IBE is its ability to be used in a re-encryption proxy system and blockchain technology.

Thus, by reducing the calculations of bilinear pairings in the phases of encryption and decryption, the amount of computations is significantly reduced compared to previous methods; as a result, the computational overhead of the system is reduced.

System Setup Phase: First, the algorithm $\mathcal{g}(1^n)$ is executed to obtain a tuple of (G, G_T, e). In this system, g is the generator element for the bilinear group G modulus p, and the identity ID is assumed to be an element of the group Z^*_p. The ciphertext is also considered an element in the group G_T. To generate public parameters and the master key, the key generation center chooses two random numbers $a_1, a_2 \in Z^*_p$. The public parameters of the system and the master key are then defined by the key generation center as Eq. (1).

$$\begin{aligned} g_1 &= g^{a_1} \\ g_2 &= g^{a_2} \\ Vo &= e(g, g_2) \\ params &= \{g, g_1, g_2, Vo, G, G_T, e, h\} \\ msk &= \{a_1, a_2\} \end{aligned} \quad (1)$$

Key Generation Phase: With the public parameters $(params)$, master key msk, and identity ID, the key generation center selects a random number r and generates the private key as Eq. (2).

$$d_{ID} = (d_1, d_2, d_3) = \left((a_1 ID + a_2)^{-1}, {g_1}^{ID^r}, {g_1}^r\right) \tag{2}$$

Encryption Phase: A random number s is selected to encrypt plaintext $m \in G_1$ using identifier $ID \in Z_p^*$. The ciphertext C_T is calculated as Eq. (3).

$$C_T = (C_1, C_2, C_3) = \left(({g_1}^{ID} g_2)^s, mVo^s, g^{ID^s}\right) \tag{3}$$

Decryption Phase: The end-user decrypts the ciphertext.$C_T = (C_1, C_2, C_3)$ using her private key and public parameters and does the Eq. (4).

$$M = \frac{C_2 e(C_3, d_3)}{e\left({C_1}^{d_1}, g_2 d_2\right)} \tag{4}$$

The correctness of the above condition is as Eq. (5).

$$M = \frac{C_2 e(C_3, d_3)}{e\left({C_1}^{d_1}, g_2 d_2\right)} = \frac{mVo^s e\left(g^{ID^s}, {g_1}^r\right)}{e\left(({g_1}^{ID} g_2)^{s^{(a_1 ID + a_2)^{-1}}}, g_2 {g_1}^{ID^r}\right)} = \frac{me(g, g_2)^s}{e(g^s, g_2)} = M \tag{5}$$

4.2. Proposed Re-Encryption Proxy

This section describes how to apply the proposed IBE to the re-encryption proxy system for use in blockchain technology.

System setup: First, the algorithm $\mathscr{g}(1^n)$ is run to obtain tuples (G, G_T, e). The generator g is then generated for the bilinear group G modulus p. Next, a one-time digital signature δ and symmetric encryption element SE are chosen. Also, three hash functions, $G: \{0,1\}^* \rightarrow Z_p^*$. $H_1: S \rightarrow G$ and $H_2: G_T \rightarrow K$ are selected in which K is the SE's keyspace. Here the ciphertext is assumed to be an element in group G_T. Random numbers $a_1, a_2, f, f' \in Z_p^*$ are selected to generate public parameters $params$ and the master key msk. The public parameters of the system and the master key are then defined by the key generation center as Eq. (6).

$$\begin{aligned} g_1 &= g^{a_1} \\ g_2 &= g^{a_2} \\ Vo &= e(g, g_2) \\ \mathcal{A} &= g^f \\ params &= \{g, g_1, g_2, Vo, G, G_T, e, h\} \\ msk &= \{a_1, a_2, f, f'\} \end{aligned} \tag{6}$$

Key Generation: With the public parameters $(params)$, the master secret key msk, and the identity ID, the key generation center selects the random numbers $r, r', n, n', z, \mu \in Z_p^*$ and generates the users' private key sk_{ID} as Eq. (7).

$$\begin{aligned} sk_{ID} &= (d_{ID}^A, d_{ID}^B, d_{ID}^C) \\ d_{ID}^A &= (d_1, d_2, d_3, d_4, d_5, d_6) \\ &= \begin{pmatrix} (a_1 ID + a_2)^{-1}, {g_1}^{ID^r}, g^{ID^r}, \mu + fr, \\ {g_1}^n, \mathcal{A}^r ({g_1}^{ID} g_2)^{-n} g^{\mu} \end{pmatrix} \\ d_{ID}^B &= (d'_1, d'_2, d'_3) \\ &= \left(\frac{fr'}{a_1 ID + a_2}, {g_1}^{n'} g^{f'ID}, \mathcal{A}^{r'} ({g_1}^{ID} g_2)^{-n'}\right) \\ d_{ID}^C &= (d_7, d_8) \\ &= {g_2}^{{a_1}^{-1}} ({g_1}^{ID} g_2)^{zID}, {g_1}^{zID} g^{f'ID} \end{aligned} \tag{7}$$

Encryption and Signature: To encrypt the plaintext $m \in G_T$ using the $ID \in Z_p^*$, a random number s and an on-time digital signature in time are selected using the private key ssk and public key svk, and the ciphertext C_{ID} is obtained as Eq. (8).

$$\begin{aligned} C_{ID} &= (C_1, C_2, C_3, C_4, C_5, C_6) \\ &= \begin{pmatrix} ({g_1}^{ID} g_2)^s, SE.Enc(h_2(V_o)^s, m), \\ {g_1}^s, h_1(svk)^s, svk, \sigma \end{pmatrix} \\ \sigma &= \delta.sig(ssk, C_1, C_2, C_3, C_4, C_5) \end{aligned} \tag{8}$$

Message Validation: The Eq. (9) verification is checked to ensure the message's validity using the ciphertext $C_{ID} = (C_1, C_2, C_3, C_4, C_5, C_6)_{ID}$, and public parameters.

$$\delta.Verify(C_5, C_6) = Yes\ , e(g_1, C_4) = e\left(C_3, h_1(C_5)\right) \tag{9}$$

Re-Key Generation: The first recipient with the identity ID is the only person who has access to the information encrypted with his ID. Upon receipt of the decryption request from the second receiver with the ID', the first receiver generates the re-encryption key for the proxy; The proxy encrypts the information with the re-encryption key without being informed of the ciphertext content. New encrypted data can be decrypted once and only by the decryption key of the second user.

In the above phases, two points are essential: first, the re-encryption key $rk_{ID \rightarrow ID'}$ is sent to the re-encryption proxy using IBE in a provided platform; Second, the proxy re-encrypts the information that it is not aware of its contents, and the risk of a re-encryption key alone does not compromise information security.

In addition, since the re-encryption key is only intended for re-encryption and does not apply to the unencrypted message, the attacker can not achieve the message content only by the re-encryption key and without the private key of the message's receiver. To generate a re-encryption key, the recipient, after selecting a random number k generates the re-encryption key $rk_{ID\to ID'}$ as Eq. (10) and provides it to the re-encryption proxy.

$$
\begin{aligned}
&rk_{ID\to ID'} = (rk_1, rk_2, rk_3)\\
&rk_1 = \frac{1}{k}\left((d_1 d_4)ID' + d'_1\right)\\
&= \frac{\mu ID' + f(rID' + r')}{k(a_1 ID + a_2)}\\
&rk_2 = {d_5}^{ID'} d'_2 = g^{f'ID} {g_1}^{(nID'+n')}\\
&rk_3 = {d_6}^{ID'} d'_3 = \frac{g^{f(rID'+r')} g^{f\mu ID'}}{{g_1}^{ID} {g_2}^{(nID'+n')}}\\
&rk_4 = {g_1}^{k} \qquad (10)
\end{aligned}
$$

Re-Encryption: The re-encryption proxy with the option of ciphertext $C_{ID} = (C_1, C_2, C_3, C_4, C_5, C_6)_{ID}$ and public parameters, re-encrypt the ciphertext as Eq. (11).

$$
\begin{aligned}
&\hat{C}_{ID} = (C'_1, C'_2, C'_3, C'_4, C'_5, C'_6)\\
&= \left(C_1, C_2, C_3, e\left({C_1}^{rk_1}, rk_4\right), rk_2, rk_3\right) \qquad (11)
\end{aligned}
$$

Decryption 2: The recipient of the message will decrypt the message by having the ciphertext $C_{ID} = (C_1, C_2, C_3, C_4, C_5, C_6)_{ID}$ that is encrypted using his ID and the private key $sk_{ID} = (d_{ID}^{A}, d_{ID}^{B}, d_{ID}^{C})$ that is in it $sk_{ID}^{A} = (d_1, d_2, d_3, d_4, d_5, d_6)$, and the public parameters are as Eq. (12).

$$
\begin{aligned}
&K = h_2\left(\frac{e({C_1}^{d_1}, g_2 d_2)}{e(C_3, d_3)}\right)\\
&SE.Dec(K, C_2) = m. \qquad (12)
\end{aligned}
$$

Finally, using the properties of element SE, plaintext m is verified.

Decryption 1: In this phase, user B, with the public parameters and the private key $d_{ID}^{C} = (d_7, d_8)$, decrypts the re-encrypted ciphertext $\hat{C}_{ID} = (C'_1, C'_2, C'_3, C'_4, C'_5, C'_6)$ as Eq. (13).

$$
\begin{aligned}
&K = h_2\left(\frac{e(C'_1, C'_5)\, e(C'_3, C'_6)\, e(C'_1, d_7)}{C'_4\, e(C'_1, d_8)}\right)\\
&SE.Dec(K, C_2) = m. \qquad (13)
\end{aligned}
$$

4.3. Proposed Blockchain-Based E-Health

As shown in Figure 1, five entities play a role in the proposed e-health system, including patients, the medical team, the key generation center, the re-encryption proxy, and the health service provider. Using a re-encryption proxy in blockchain technology and combining it with the proposed encryption scheme in the cloud health system, we have the following scenario (It should be noted that all steps in this scenario are performed by Section 5.2):

0. The key generation center generates the public parameters of the system and the public and private keys of each entity and provides them. This process takes place after registering the identity of each entity in the health system and authenticating them.
1. Patient A visits the medical team for the first time. The medical team sends their ID (the public key) to the patient so that both can use identity-based encryption and agree on a symmetric key to encrypt the patient's medical records.
2. The patient encrypts his electronic health documents to provide confidentiality using the agreed symmetric key and AES encryption (or similar). Then, according to the system model, it outsources these medical documents to the cloud, along with the symmetric agreement key encrypted in identity-based encryption.
3. Members of a blockchain network extract encrypted data from the cloud and execute a verification algorithm to message validation. Then, by performing a consensus protocol, they select the bookkeeping member.
4. The bookkeeping member sends the encrypted data to the cloud re-encryption proxy and obtains the data access address from the cloud.

4.1. It also writes a brief description that includes encrypted data and its cloud address as a specific format in the blockchain.

5. Whenever the medical team needs patient A's electronic health records, they review the blockchain's contents and obtain the data address they want.

5.1. The medical team receives the relevant encrypted documents from the cloud and, with a pre-shared symmetric key, decrypts the documents easily.

6. We assume that user A becomes ill again sometime later and needs the services of another medical center. The health service provider needs the information of user A to diagnose the disease better and take action. Therefore, they find the address of information of patient A in the blockchain.

6.1. The health service provider sends the address of information of patient A in the blockchain and their ID to the medical team.

7. The medical team, having its private key and the ID of the health service provider, generates a re-encryption key by running the re-key generation algorithm in the re-encryption proxy system. It also provides the address of information of patient A and the re-encryption key to the re-encryption proxy.

7.1. In addition, it sends the decryption key of the first-level encrypted text to the requester for information on user A.

8. The re-encryption proxy re-encrypts the pre-shared symmetric key that has already been encrypted using the identity-based encryption scheme and sends it to the health service provider.
9. The health service provider decrypts patient A's medical information using the key provided to them by the medical team during the re-key generation key phase and the pre-shared symmetric key that has been re-encrypted by the proxy and sent to themThis section describes how to apply the proposed IBE to the re-encryption proxy system for use in blockchain technology.

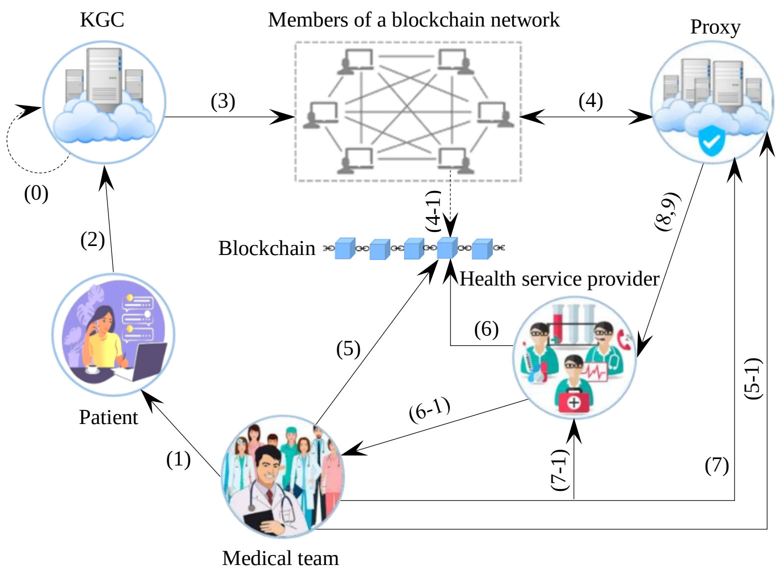


Figure 1. The framework of the proposed e-health system

5. The Proposed Scheme Analysis

In this section, the proposed identity-based encryption scheme is evaluated in terms of security and performance. First, the security of the proposed scheme is proved in the standard model, and it is shown that this method has provable security in the random oracle model. Then, the efficiency of the proposed scheme is compared with other identity-based encryption methods.

5.1. Security Analysis

The IND-ID-CCA security model is used here to prove the proposed IBE formally. This security model is based on a game between the challenger and the attacker in 5 steps (setup, phase 1, challenge, phase 2, and guess). According to theorm1, the proposed method is secure against the IND-ID-CCA attacker. In the following, we explain the mentioned steps.

- **Theorm1:** If the hypotheses of Bilinear Decisional Diffie-Hellman (BDDH) are valid in the tuple (G, G_T, e), the proposed identity-based encryption scheme is secure against the IND-ID-CCA attacker.
- **Proof:** Suppose attacker α can attack the proposed scheme. In this case, we show that a structure called Algorithm β can solve the BDDH problem in the tuple (G, G_T, e). Algorithm β has multiple (g, g^a, g^b, g^c, T). T is a random number or equal to $e(g,g)^{acb}$. The purpose of Algorithm β is to generate the bit 1 if the equation $T = e(g,g)^{abc}$ is satisfied; If T equals a random number, the bit 0 is generated as the output. Suppose $a = a_1$ and $b = a_2$. Algorithm β interacts with attacker α to solve the problem ahead. The game of the selective identity begins with the first output of the identity ID^* that attacker α intends to attack as the following:

- **Setup:** To generate system parameters, Algorithm β selects two random numbers, a and b, and generates the public parameters $params = \{g, g_1, g_2, Vo\}$, and passes them to attacker α.
- **Phase 1:** In this phase, the attacker α selects identities adaptively. It then uses these identities to issue private key generation or decryption requests. Algorithm β provides a valid decrypted text related to the attacker's requested ciphertext and a private key associated with the attacker's desired ID.
- **Challenge:** After the attacker α decides to end phase one, it sends two plaintexts m_1, m_2 with equal length and an identity ID_* to the algorithm β. Algorithm β then selects the random bit $\gamma \in \{0, 1\}$ and generates the ciphertext $CT_* = (g_1^{ID_*} g_2)^c, M \cdot T, g^{{ID_*}^c}$. It then sends the ciphertext CT_* as a challenge to attacker α. Note that algorithm β does not know the value of c; But receiving the value of g^c is considered as one of the inputs to the BDDH problem.

- **Phase 2:** This phase is similar to phase 1; the difference is that the attacker cannot request a private key for the identity ID^* and a request for decryption CT_*.
- **Guess:** In the final, attacker α guesses the value $\gamma' = \{0, 1\}$. If the equation $\gamma' = \gamma$ is true, Algorithm β displays the value 1 as the output. This means $T = e(g, g)^{acb}$; Otherwise, it represents the value 0, which means T is a random number. If the equation $T = e(g, g)^{acb}$ is satisfied, attacker α can defeat the proposed scheme, which means that algorithm β must exist and solve the hard problem of BDDH. As a result, as long as no algorithm β can solve the hard BDDH problem, the proposed encryption scheme will be IND-ID-CCA secure.

5.2. Security Comparison

Table 2 shows a comparison between the proposed method and other methods from a security point of view. Protocols that benefit from encryption, digital signature, re-encryption proxy, and blockchain are more secure. Also, IBE methods whose security has been proven in the standard model are more reliable.

Table 2: Security comparison

Schemes	Security Properties							
	Confidentiality	Traceability	Authentication	Non-repudiation	Integrity	Optimized in re-encryption proxy	Access control with blockchain	Proof security in the standard model
Boneh and Franklin [4]	✓	✓	×	×	×	×	×	×
Boneh and Franklin [22]	✓	✓	×	×	×	×	×	×
Sakai and Kasahara [23]	✓	✓	✓	✓	✓	×	×	×
Boneh and Boyen [24]	✓	✓	×	×	×	×	×	✓
Gentry [25]	✓	✓	×	×	×	×	×	✓
Boneh and Boyen [26]	✓	✓	×	×	×	×	×	✓
Bakhtiari and Hosseinzadeh [30]	✓	✓	✓	✓	✓	×	×	×
Wang et al. [3]	✓	✓	✓	✓	✓	✓	×	✓
Wang and Song [7]	✓	✓	✓	✓	✓	×	✓	✓
Proposed scheme	✓	✓	✓	✓	✓	✓	✓	✓

5.3. Performance Evaluation

Less use of bilinear pairing in identity-based encryption schemes reduces computational overhead. Therefore, Table 3 shows the number of bilinear pairing functions used in the three phases of key generation, encryption, decryption in IBE schemes, and their comparison with the proposed method.

Table 3: Efficiency comparison

Schemes	Need to bilinear pairing		
	KGC	Encryption	Decryption
Boneh and Franklin [4]	×	✓	✓
Boneh and Franklin [22]	×	✓	✓
Sakai and Kasahara [23]	✓	×	✓
Boneh and Boyen [24]	✓	×	✓
Gentry [25]	✓	×	✓
Boneh and Boyen [26]	✓	×	✓
Bakhtiari and Hosseinzadeh [30]	×	✓	✓
Wang et al. [3]	×	✓	✓
Wang and Song [7]	✓	✓	✓
Proposed scheme	✓	×	✓

Then, the superior methods with the least use of bilinear pairings are compared with the proposed method in running time. We obtained the listed results by implementing the mentioned schemes using the JPBC library in the programming language of Java [37] and an environment with specifications stated in Table 4.

Table 4: System specifications for implementation

Platform	Specifications
Integrated development environment	Eclipse
The operating system	Ubuntu 16.04
CPU	Intel Core i7 3.5GHz
RAM	8GB

Figure 2 shows the running time of the setup and key generation phases of the considered methods in milliseconds. This type is known as type A in the JPBC library. This part of the system, which includes generating public parameters and the master key, is calculated only once in a cryptographic method. The pre-calculated parameters are used in the setup phase in the next times of the key generation, encryption, and decryption phases.

Figure 3 shows the encryption and decryption running time in identity-based encryption schemes. Some identity-based encryption methods also use the bilinear pairing function in the encryption phase. In this case, with each running of the encryption phase,

there is a need to calculate bilinear pairing functions. This increases the cost of an IBE system. For this reason, IBE methods that use bilinear pairing calculations only in the decryption phase significantly reduce the computational cost.
As shown in Figure 2, the proposed IBE has a similar running time in setup and key generation phases compared to the method of Boneh and Boyen [24]. Figure 3 also shows that the proposed method in the encryption phase has a 13-millisecond advantage over Boneh and Boyen [24]; But in the decryption phase, the method of Boneh and Boyen [24] has a 3-millisecond advantage over the proposed method.
In this way, in total, the running time of encryption and decryption of the proposed IBE has a shorter running time (about 10 milliseconds), which reduces the cost of the system.
In addition, the proposed IBE method, compared to the IBE method of Wang et al. [3], shows significant changes in reducing the running time of each of the three phases of key generation, encryption, and decryption. Figure 4 compares the running time of the proposed re-encryption proxy with the re-encryption proxy Wang et al. [3]. As the implementation results show, the proposed method reduces the cost of the re-encryption proxy scheme by reducing the computational overhead and improving the running time. In addition, the proposed scheme offers high flexibility and scalability due to the use of blockchain technology.

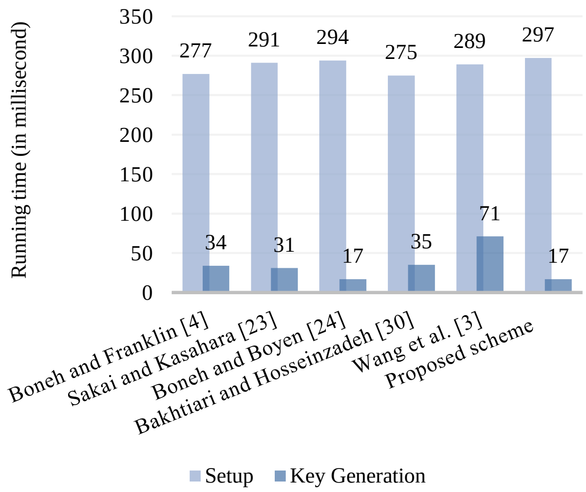


Figure 2. Compare of running times of setup and key generation phases

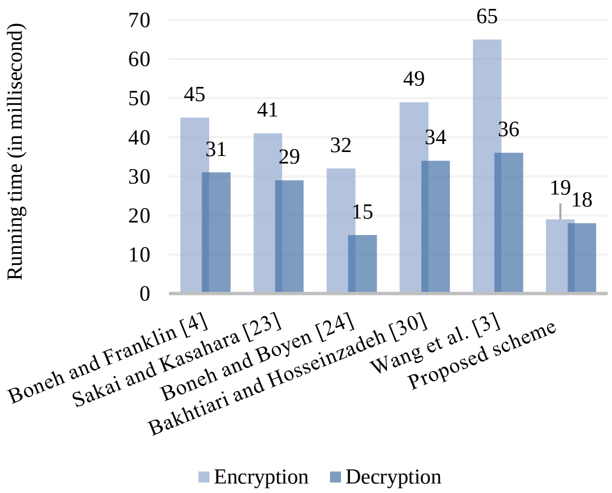


Figure 3. Compare of running times of encryption and decryption phases

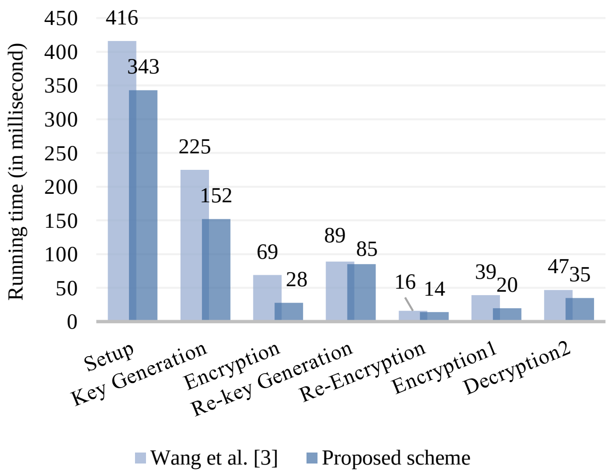


Figure 4. Compare of running times of the re-encryption proxy

6. Conclusion

In this paper, a secure and efficient cloud health system is proposed for sharing medical data between patients, hospitals, and other healthcare centers. The system includes a new identity-based encryption scheme with improvements in the re-encryption proxy. It can also be used in blockchain technology.
In the proposed method, the content of medical records is stored in the cloud servers, and their access address is located in the blockchain. This work has led to improved memory usage of blocks and more efficient medical data management during sharing. In addition, patients' medical records are encrypted and provided to the members of a blockchain network and the proxy in the cloud; No entity other than authorized users will be able to access the content of patients' medical information.
Also, in the proposed identity-based encryption scheme, bilinear pairing is used only in the decryption

phase, and the pairing parameter used in the encryption phase is a predetermined parameter. Another strength of the proposed approach is shortening the sending parameters while preserving system security.

The security analysis showed that the proposed scheme provides more security properties than similar schemes. In this regard, we also proved the security of the proposed IBE in the standard model. In addition, the implementation results showed that the proposed scheme has better running time and lower computational cost than previous similar schemes.

Acknowledgment

The authors sincerely thank this journal for giving chances to proposing the scheme.

References

[1] H. Toral-Cruz, D. He, A.D. Mihovska, K.K. Raymond Choo, and M.K. Khan, Reliable and Secure e-Health Networks. Wireless Personal Communications, (2021), 117, 1–6. https://doi.org/10.1007/s11277-021-08104-z.

[2] W. Hsin-Te, and T. Chun-Wei, Toward Blockchains for Health-Care Systems: Applying the Bilinear Pairing Technology to Ensure Privacy Protection and Accuracy in Data Sharing. IEEE Consumer Electronics Magazine, 7 (2018), 65-71. https://doi.org/10.1109/MCE.2018.2816306.

[3] X.A. Wang, J. Ma, F. Xhafa, M. Zhang, and X. Luo, Cost-Effective, Secure E-Health Cloud System Using Identity-Based Cryptographic Techniques. Future Generation Computer Systems, 67 (2017), 242–254. https://doi.org/10.1016/j.future.2016.08.008.

[4] D. Boneh, and M. Franklin, (2001, August). Identity-Based Encryption from the Weil Pairing. In Annual International Cryptology Conference (pp. 213–229). Springer, Berlin. https://doi.org/10.1007/3-540-44647-8_13.

[5] P. Patil, M. Sangeetha, and V. Bhaskar, Blockchain for IoT Access Control, Security and Privacy: A Review. Wireless Personal Communications, 117 (2021), 1815–1834. https://doi.org/10.1007/s11277-020-07947-2.

[6] P. Pandey and R. Litoriya, Securing E-health Networks from Counterfeit Medicine Penetration Using Blockchain. Wireless Personal Communications, 117 (2021), 7–25. https://doi.org/10.1007/s11277-020-07041-7.

[7] H. Wang and Y.J. Song, Secure Cloud-Based EHR System Using Attribute-Based Cryptosystem and Blockchain. Journal of Medical Systems, 42(8) (2018), 152:1–152:9. https://doi.org/10.1007/s10916-018-0994-6.

[8] V. Della Mea, What Is E-Health (2): The Death of Telemedicine? Journal of Medical Internet Research, 3(2) (2001). https://doi.org/10.2196/jmir.3.2.e22.

[9] O. Enaizan, A.A. Zaidan, N.H.M. Alwi, B.B. Zaidan, M.A. Alsalem, O.S. Albahri, and A.S. Albahri, Electronic Medical Record Systems: Decision Support Examination Framework for Individual, Security and Privacy Concerns Using Multi-Perspective Analysis. Health and Technology, 10 (2020), 795–822. https://doi.org/10.1007/s12553-018-0278-7.

[10] J. O'donoghue, and J. Herbert, Data Management within M-Health Environments: Patient Sensors, Mobile Devices, and Databases. Journal of Data and Information Quality, 4 (2012), 1–20. https://doi.org/10.1145/2378016.2378021.

[11] L. Neuhauser, and G.L. Kreps, Online Cancer Communication: Meeting the Literacy, Cultural and Linguistic Needs of Diverse Audiences. Patient Education and Counseling, 71 (2008), 365–377. https://doi.org/10.1016/j.pec.2008.02.015.

[12] L. Neuhauser, and G.L. Kreps, E-Health Communication and Behavior Change: Promise and Performance. Social Semiotics, 20 (2010), 9–27. https://doi.org/10.1080/10350330903438386.

[13] G.L. Kreps, Strategic Use of Communication to Market Cancer Prevention and Control to Vulnerable Populations. Health Marketing Quarterly, 25 (2008), 204–216. https://doi.org/10. 1080/07359680802126327.

[14] J. Benaloh, M. Chase, E. Horvitz and K. Lauter, Patient-Controlled Encryption: Ensuring the Privacy of Electronic Medical Records. In Proceedings of the 2009 ACM Workshop on Cloud Computing Security 2009, 103–114. https://doi.org/10.1145/1655008.1655024.

[15] Y. Xue, X. Mao, Y. Guo, and S. Lv,The Research Advance of Facial Expression Recognition in Human-Computer Interaction. Journal of Image and Graphics, 5 (2009), 764–772.

[16] M. Barua, X. Liang, R. Lu and X. Shen, ESPAC: Enabling Security and Patient-Centric Access Control for E-Health in Cloud Computing, International Journal of Security and Networks, 6 (2011), 67–76. https://doi.org/10.1504/IJSN.2011.043666.

[17] L. Guo, C. Zhang, J. Sun and Y. Fang, Paas: A Privacy-Preserving Attribute-Based Authentication System for E-Health Networks. IEEE Transactions on Mobile Computing, 13 (2012), 1927-1941. https://doi.org/10.1109/ICDCS.2012.45.

[18] M. Li, S. Yu, Y. Zheng, K. Ren and W. Lou, Scalable and Secure Sharing of Personal Health Records in Cloud Computing Using Attribute-Based Encryption. IEEE Transactions on Parallel and Distributed Systems, 24 (2013), 131–143. https://doi.org/10.1109/TPDS.2012.97.

[19] H. Yan, J. Li, X. Li, G. Zhao, S.Y. Lee and J. Shen, Secure Access Control of E-Health System with Attribute-Based Encryption. Intelligent Automation & Soft Computing, 22 (2016), 345–352. https://doi.org/10.1080/10798587.2015.1132586.

[20] X. Boyen, A Tapestry of Identity-Based Encryption: Practical Frameworks Compared. International Journal of Applied Cryptography, 1 (2008), 3–21. https://doi.org/10.1504/IJAC T.2008.017047.

[21] A. Shamir, Identity-Based Cryptosystems and Signature Schemes. Lecture Notes in Computer Science, 84 (1984), 47–53. https://doi.org/10.1007/3-540-39568-7_5.

[22] D. Boneh and M. Franklin, Identity-Based Encryption from the Weil Pairing. SIAM Journal on Computing, 32(3) (2003), 586–615. https://doi.org/10.1137/S0097539701398521.

[23] R. Sakai and M. Kasahara, ID-Based Cryptosystems with Pairing on Elliptic Curve. IACR Cryptology ePrint Archive. 2003.

[24] D. Boneh and X. Boyen, Efficient Selective-ID Secure Identity-Based Encryption without Random Oracles. In International Conference on the Theory and Applications of Cryptographic Techniques, 2004, 223–238. https://doi.org/10.100 7/978-3-540-24676-3_14.

[25] C. Gentry, Practical Identity-Based Encryption without Random Oracles. In Annual International Conference on the Theory and Applications of Cryptographic Techniques, 2006, 445–464. https://doi.org/10.1007/11761679_27.

[26] D. Boneh and X. Boyen, Efficient Selective Identity-Based Encryption without Random Oracles. Journal of Cryptology, 24 (2011), 659–693. https://doi.org/10.1007/s00145-010-9078-6.

[27] D. Galindo, Chosen-Ciphertext Secure Identity-Based Encryption from Computational Bilinear Diffie-Hellman. In International Conference on Pairing-Based Cryptography, 2010, 367–376. https://doi.org/10.1007/978-3-642-17455-1_23.

[28] J.H. Park, K. Lee and D.H. Lee, New Chosen-Ciphertext Secure Identity-Based Encryption with a Tight Security Reduction to the Bilinear Diffie-Hellman Problem. Information Sciences, 325 (2015), 256–270. https://doi.org/10.1016/j.in s.2015.07.011.

[29] W. Susilo, F. Guo and Y. Mu, Efficient Dynamic Threshold Identity-Based Encryption with Constant-Size Ciphertext. Theoretical Computer Science, 609 (2015), 49–59. https://doi.org/10.1016/j.tcs.2015.09.006.

[30] S. Bakhtiari-Chehelcheshmeh and M. Hosseinzadeh,A New Certificateless and Secure Authentication Scheme for Ad-Hoc Networks. Wireless Personal Communications, 94 (2017), 2833–2851. https://doi.org/10.1007/s11277-016-3721-y.

[31] M. Green and G. Ateniese, Identity-Based Proxy Re-Encryption. In International Conference on Applied Cryptography and Network Security, 2007, 288–306. https://doi.org/10.100 7/978-3-540-72738-5_19.

[32] S. Nakamoto, Bitcoin: A Peer-to-Peer Electronic Cash System,2012.

[33] N. Kshetri, Blockchain's Roles in Strengthening Cybersecurity and Protecting Privacy. Telecommunications Policy, 41 (2017), 1027-1038. https://doi.org/10.1016/j.telpol.2017.09.003.

[34] X. Yue, H. Wang, D. Jin, M. Li and W. Jiang, Healthcare Data Gateways: Found Healthcare Intelligence on the Blockchain with Novel Privacy Risk Control. Journal of Medical Systems, 40 (2016), 218. https://doi.org/10.1007/s10916-016-0574-6.

[35] M. Banerjee, J. Lee and K.K. Raymond Choo,A Blockchain Future for the Internet of Things Security: A Position Paper. Digital Communications and Networks, 4 (2018), 149-160. https://doi.org/10.1016/j.dcan.2017.10.006.

[36] A. Azaria, A. Ekblaw, T. Vieira and A. Lippman, Medrec: Using Blockchain for Medical Data Access and Permission Management. Open and Big Data (OBD), International Conference on IEEE, 2016, 25-30. https://doi.org/10.1109/OB D.2016.11.

[37] A. De Caro, V. Iovino and A. Renato, JPBC : Java Pairing Based Cryptography. IEEE Symposium on Computers and Communications, 2011, 850–855. https://doi.org/10.1109/IS CC.2011.5983948E.

Journal of Optimization of Soft Computing (JOSC)

Journal of Optimization of Soft Computing (JOSC)

Vol. 1, Issue 1, pp: (12-22), September-2023

Journal homepage: https://sanad.iau.ir/journal/josc

Extending the Lifetime of Wireless Sensor Networks Using Fuzzy Clustering Algorithm Based on Trust Model

Farshad Kiyoumarsi[1*], Behzad Zamani Dehkordi[2]

[1,2]Department of Engineering, Faculty of Computer, Shahrekord Branch, Islamic Azad University, Shahrekord,Iran

Article Info

Article History:
Received :2023/7/4
Revised :2023/7/20
Accepted 2023/7/23

Keywords:
Cluster head, Fuzzy clustering, Network lifetime, Trust model, Wireless sensor network.

**Corresponding Author's Email Address:* farshad.kiyoumarsi@iau.ac.ir

Abstract

Wireless sensor networks (WSNs) are the safest and most widely used existing networks, which are used for monitoring and controlling the environment and obtaining environmental information in order to make appropriate decisions in different environments. One of the very important features of wireless sensor networks is their lifetime. Two important factors come to mind to increase the lifetime of networks: These factors are maintaining the coverage of the network and reducing the energy consumption of sensor nodes simultaneously with the uniform consumption of energy by all of them. Clustering, as the optimal method of data collection, is used to reduce energy consumption and maintain the coverage of the network in wireless sensor networks. In clustered networks, each node transmits acquired data to the cluster head to which it belongs. After a cluster head collects all the data from all member nodes, it transmits the data to the base station (sink). Given that fuzzy logic is a good alternative for complex mathematical systems, in this study, a fuzzy logic-based trust model uses the clustering method in wireless sensor networks. In this way, cluster-head sensors are elected from among sensors with high reliability with the help of fuzzy rules. As a result, the best and most trusted sensors will be selected as the cluster heads. The simulation results in MATLAB software show that in this way, in comparison with K-Means, FCM, subtractive clustering, and multi-objective fuzzy clustering protocols, the energy consumption in clustered nodes will decrease and the network's lifetime will increase.

1. Introduction

One of the most difficult topics in computer and electronics sciences today is the discussion of remote monitoring and controlling systems [1]. Production and deployment of miniature, battery-powered nodes that communicate through wireless links have been made possible by earlier developments in information technology (IT), particularly in MEMS (micro electro-mechanical systems). A single node can collect information from the area that is in its coverage area. These nodes collaborate in order to have conceptually significant information from the whole area. Wireless sensor networks are made up of such nodes having sensing capabilities. Early implementation aimed to passively utilize these sensors for indoor applications. Scalar information like as temperature, humidity, pressure, and the location of nearby objects can all be sensed by these early nodes. These nodes are initially underpowered in terms of compute and storage, and their sole purpose is to send scalar data to the base station (sink). However, new sensor nodes outperform their predecessors in terms of compute power, storage capacity, and power management, and their main application domain switches from indoor to outdoor applications. Researchers are quite interested in the energy strategies of sensor nodes because they typically have batteries that cannot be recharged. For all of these reasons, one of the main objectives is still to reduce energy use through energy efficiency. To increase the lifespan of sensor nodes in this regard, energy-efficient algorithm design is essential [2].

Sensor nodes in the WSN can be organised into discrete groups known as clusters. A cluster-head (CH), also known as the leader, controls data aggregation from member nodes and transfer of the compiled data to the sink in each cluster. CH selection may be carried out centrally or decentralised. With a lot of sensor nodes, clustering in WSNs ensures strict performance requirements [3, 4]. Additionally, it makes WSNs more scalable [5]. Other benefits of clustering include route setup localisation, communication bandwidth conservation by minimising relayed packets, a decrease in the rate of energy consumption, and network topology stabilisation [6]. The literature has extensively examined the selection mechanisms since effective CH selection can lower energy consumption. The majority of strategies use a two-stage procedure, choosing CHs with more energy left over in the first step, and rotating the member nodes in the second step to balance energy consumption. This example demonstrates how these selection methods solely consider the nodes' energy and disregard their trust. Because the trust of the deployed nodes is not taken into account, clusters that are close to the sink are generated in lower sizes than clusters that are further away.

This study introduces a new clustering approach with the aim of prolonging the lifetime of WSNs, which is not only energy-efficient but also distribution-independent for wireless sensor networks. This protocol determines the radius of the sensors with the help of fuzzy logic and considering distance and energy variables, then determines the chance of becoming CH by considering the energy of the sensors and the number of neighboring sensors in the determined radius, and finally chooses sensors as cluster heads, which have high trust and a high chance of becoming CH.

The remainder of this paper is organized as follows: In Section 2, related studies are summarized. The system model is given in Section 3. Our proposed clustering algorithm, FCATM, is introduced and discussed in Section 4. Thereafter, simulation results and performance evaluations are explored in Section 5. Finally, in Section 6, our conclusions and possible future works are given.

2. **Related Works**

Effective data collection from deployed nodes is the focus of the data aggregation procedure. In this regard, clustering techniques offer energy-efficient infrastructure for the required activity. The established needs, such as reducing the quantity and size of data packets to be conveyed and offering effective delivery mechanisms for these routed packets, lead to the need for clustering. When taking into account the application kinds, which involve more multimedia streaming data every other day, this subject becomes even more important. Several WSN clustering techniques have been suggested in the literature. Key and differentiating characteristics of the popular clustering algorithms are described in the paragraphs that follow. It is helpful to think about what other existing clustering algorithms perform in order to assist identify the important elements of our suggested approach [4].

LEACH is a clustering-based system that uses random base station rotation to divide the energy load across the network's sensors in an equitable manner. In order to enable scalability and robustness for dynamic networks, LEACH employs localised coordination. Additionally, it adds data fusion into the routing protocol to lessen the quantity of data that needs to be broadcast to the base station. The probability that each node will become a cluster head in this round is determined by a stochastic method at the beginning of each round; in other words, this protocol chooses CHs using a probability model and then rotates CHs. Additionally, neither the gateway nor the nodes negotiate the CHs to be chosen in the following round. Before sending data to the sink in LEACH, the CHs compress the data. LEACH does not take into account the distribution of sensor nodes or the remaining energy on each node, hence it is not a lifetime-efficient algorithm for the network. [7].

Node equality is the main presumption of the Hybrid Energy Efficient and Distributed (HEED) algorithm, which is created for multi-hop networks [8]. It chooses cluster heads on a periodic basis using a hybrid of two node parameters: first cluster heads are chosen based on residual energy, and final cluster heads are chosen based on intra-cluster communication costs. However, because of its propensity to produce more clusters than predicted, the HEED algorithm suffers from the hotspots problem and results in uneven energy usage. [9].

Due to its effectiveness and simplicity in grouping huge data sets, the K-Means algorithm is one of the most often used clustering algorithms [10]. In the conventional K-Means technique, a collection of data, set D, is categorized using a set of initialized apriority clusters (k clusters). It first defines k centroids, one for each cluster, after which it takes into account data objects from the given data set and links them to the closest centroid. The distance between data objects and the centroids is often calculated using the Euclidean distance. When early grouping is complete and there are no more data objects, the first stage is finished. Here, new centroids must be calculated from scratch. The same data objects are bound to the nearest centroid and create a loop after acquiring fresh centroids. K-centroids gradually shift their points at the conclusion of the loop until they stop moving altogether [10, 11, 19]. The foundation of this approach is the squared error function minimization. The K-Means algorithm has issues with providing an initial seed value and a preliminary number of clusters. This approach also depends on the original cluster seed values and always converges to a local minimum. [10].

Fuzzy logic is being used by an increasing number of clustering algorithms to solve the issues that arise as a result of the uncertainties that exist in WSN nature. They are referred to as fuzzy clustering techniques as a result. Fuzzy logic is primarily used in these methods to combine the relevant input factors in a better way to produce the desired output, which in this case is CH election. [12, 18].

The data to be analyzed must be in the form of numerical vectors called feature vectors, and the number of clusters must be predefined in order to acquire the membership values of the feature vectors, according to Bezdek [13, 20]. The fuzzy c-means clustering algorithm's objective function that needs to be minimized is:

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}{}^{m} \| x_i - c_j \|^2 \qquad 1 \le m < \infty \tag{1}$$

Where cj is the cluster-head of this cluster and uij is the degree of the membership function of xi in cluster j. The total number of nodes is N, the total number of clusters is C, and the parameter m, which affects the fuzziness of the generated clusters, is larger than 1.

The fuzzy c-means iterative algorithm has been described by Bezdek and Pal as a numerical procedure in their classification technique. With the update of membership uij and the cj cluster centres by [14], the goal function described above is optimised iteratively to perform fuzzy partitioning.

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left[\frac{\| x_i - c_j \|}{\| x_i - c_k \|}\right]^{\frac{2}{m-1}}} \tag{2}$$

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}{}^{m} . x_i}{\sum_{i=1}^{N} u_{ij}{}^{m}} \tag{3}$$

This iteration will stop when: maxij

$$\{| u_{ij}{}^{k+1} - u_{ij}{}^{k} |\} < \varepsilon \tag{4}$$

Where ε is a termination criterion between 0 and 1, and k is the number of iteration steps..

This procedure converges to a local minimum or a saddle point of Jm.

The algorithm is composed of the following steps:

1. Initialize U = [uij] matrix, U (0)
2. At k-step: calculate the centers vectors C(k)=[cj] with U(k)

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}{}^{m}.x_i}{\sum_{i=1}^{N} u_{ij}{}^{m}} \tag{5}$$

3. Update U(k), U(k+1)

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left[\frac{\| x_i - c_j \|}{\| x_i - c_k \|}\right]^{\frac{2}{m-1}}} \tag{6}$$

If ||U(k+1) - U(k)||<ε then STOP; otherwise return to step 2.

The best data point to use to construct a cluster centroid based on the density of nearby data points is found using subtractive clustering [15]. Chiu's mountain method is expanded upon by this strategy. The mountain method for clustering is relatively straightforward and efficient. The mountain approach's accompanying computing method is solved using the subtractive-clustering algorithm. The calculation of this approach is proportional to the size of the problem and uses data points as candidates for the cluster centre. It makes an estimate regarding the cluster centres' initial size and position. The potential for each data point is calculated based on its distance from the actual data point once the data space is divided into grid points. The grid point with the highest potential value will be selected as the initial cluster centre since it will have a high potential value due to the proximity of many data points. We will attempt to locate the second cluster centre by determining the grid point with the highest potential value after choosing the first cluster centre. The next cluster centre will be a grid with numerous data points nearby in addition to the first cluster centre grid point because grid points close to the first cluster centre will lower its potential worth. Up until the potential of every grid point drops below a threshold value, this process of gaining new cluster centres and lowering the potential of neighbouring grid points is repeated. As a result, this technique is among the easiest and most efficient ways to identify cluster centres. However, the complexity of its processing exponentially increases as data dimensions increase.

Take a look at the following set of n data points: X = x1, x2, x3,..., xn. Each point is then taken into account as a potential cluster centre. The following is a definition of the data point's xn potential:

$$p_n = \sum_{i=1}^{n} e^{\frac{-4 x_n - x_i{}^2}{r_a{}^2}} \tag{7}$$

The neighbourhood is defined by the positive constant ra, where ra is the hypersphere cluster radius in the data space. Choose the data point with the highest potential after determining each data point's ability to serve as the first cluster centre. Consider x1 and p1 to be the first cluster centre and their respective potentials. Use the formula below to then update each data point's potential.

Think of a set of n data points:

X is equal to "x1, x2, x3,..., xn". Each point is then taken into account as a potential cluster centre. The following is a definition of the data point's xn potential:

$$p_n = p_n - p_1 e^{\frac{-4 x_n - x_1{}^2}{r_b{}^2}} \tag{8}$$

The positive constant rb is the hypersphere penalty radius in data space. Find the next highest potential to serve as the following cluster centre after computing the revised potential of each data point. These procedures keep going until there are enough cluster centres.

The Multi-objective Fuzzy Clustering Algorithm (MOFCA) [2] is built with two key considerations in mind: first, it must be energy efficient in all circumstances where it can be used, and second, it must be light enough to be installed on actual sensor hardware boards. It is a distributed unequal fuzzy clustering algorithm that uses local judgements to choose the tentative and final CHs and determine the node competition radius. To determine the competition radius for tentative CHs, MOFCA takes into account three variables: the distance to the sink, the node's remaining energy, and the node's density. Like several other methods, MOFCA uses fuzzy logic in addition to these factors to determine the competition radius. Compared to the protocols discussed above, this protocol's algorithm is more energy-efficient, and its performance scales well.

3. System Model

The system model, together with its network and energy model subsections, is thoroughly explained in this section.

3.1. Network Model

The following properties are assumed with regard to the sensor network being studied:

- The nodes are determined as the base station (the sink), root, and member nodes.
- All nodes are identical.
- The capabilities of all nodes, such as processing and communicating, are similar.
- All sensors are located in a two-dimensional space, and information on the location of each one, known as a basic premise.
- The nodes are deployed either manually in order to form a non-uniform or randomly.
- The base station may be located anywhere within the WSN's Area of Interest (AOI). It is not necessary for it to be far from the sensing area. But it can also be outside the AOI.
- Following the deployment phase, every sensor node must be stationary. However, the forcible modification of the initial placement by remote control is not included in the definition of "mobility" in this context. It only includes changes to locations brought on by tectonic movements like erosion or displacement brought on by outside things. The inclusion of this supposition also targets emerging networks.
- Because mobility is assumed to be generated by external sources, it does not cause nodes to consume energy.
- All sensor nodes have the same amount of energy when they are deployed.
- All sensor nodes have the same rate of production data and send information periodically.
- The base station’s power level in comparison with the energy of sensors is unlimited.
- The nodes are capable of adjusting transmission power according to the distance of the receiving nodes.
- The distance between nodes can be calculated based on radio signal strength.

3.2. Energy Model

Sensor nodes consume energy mainly during packet transmitting, packet receiving, sensing, and data processing. We used the energy model given in [2, 16]. Depleted energy measurement in transmitting or receiving l bits over a distance of d is done as in Eqs. (9) and (10), respectively. E_{elec}= 50 nJ/bit, E_{fs}= 10 pJ/bit/m2, Emp= 0.0010 pJ/bit/m4, and d0 = 20 m. E_{elec} is the energy consumption per bit in the transmitter and receiver circuitry, Emp is the energy dissipated per bit in the RF amplifier, and L is the length of packets.

$$E_{TX} = \begin{cases} L \times E_{elec} + L \times E_{fs} \times d^2 & d < d0 \\ L \times E_{elec} + L \times E_{mp} \times d^4 & d \geq d0 \end{cases} \quad (9)$$

$$E_{RX} = L \times E_{elec} \quad (10)$$

In a wireless sensor network, cluster heads are responsible for collecting data from the sensors and sending it to the base station. Therefore, the energy consumed in cluster heads during a round is calculated by the following equation:

$$E_{CH} = (\frac{n}{k} - 1) L \cdot E_{elec} + \frac{n}{k} L \cdot E_{DA} + L \cdot E_{elec} + L \cdot E_{fs} \cdot d^4_{to\text{-}BS} \quad (11)$$

Where n is the total number of nodes, k is the total number of clusters, EDA is the energy or cost of gathering relevant data from all sensors in a cluster by the cluster head, and $d_{to\text{-}BS}$ is the average distance between the cluster head and the base station.

The energy used in any typical sensor node is calculated by the following equation:

$$E_{CM} = L \cdot E_{elec} + L \cdot E_{fs} \cdot d^2_{to\text{-}CH} \quad (12)$$

$d_{to\text{-}CH}$ is the average distance between the cluster head and the typical sensor node.

4. Fuzzy Clustering Algorithm Based on Trust Model (FCATM)

In this study, it is assumed that the sensor nodes get a message from the base station and calculate their distance to the base station. Also, each sensor is aware of the situation and the extent of its residual energy.

The proposed protocol selects local cluster heads so that the base station doesn't need to gather all nodes' information and determine all cluster nodes. Furthermore, the proposed protocol is similar to the LEACH protocol, since time is divided into sections called rounds and clusters are configured in each round. In each round, there are two phases: startup and steady state.

In the startup phase, the impact radius and CH chance of the sensors are calculated. Then cluster heads are elected based on chance and the trust of their neighbours.

In the steady state phase, CH data collection, aggregation, and transmission to the base station are done.

4.1. Determining the impact radius with the help of fuzzy logic

As previously mentioned, each sensor is aware of the extent of its residual energy and its distance to the base station. FCATM considers an impact radius for each sensor node to reduce energy consumption according to residual energy and the distance to the base station. This radius limits the sensor node to its impact radius if elected as CH and prevents rapid depletion of energy. In the fuzzy function, fuzzy rules use three fuzzy variables to determine the impact radius. They are:

- Energy: the amount of remaining energy in the sensor node.
- Distance: the separation between the sensor node and the base station.
- Radius: The sensor node's impact radius.

Fuzzy rules used to calculate the impact radius of sensor nodes are shown in Table 1:

Table 1

Fuzzy rules to determine impact radius.

NO.	ENERGY	DISTANCE	RADIUS
1	Low	Far	Very Small
2	Low	Medium	Very Small
3	Low	Close	Small
4	Medium	Far	Small
5	Medium	Medium	Medium
6	Medium	Close	Large
7	High	Far	Medium
8	High	Medium	Large
9	High	Close	Very Large

4.2. Calculating the chance of becoming CH with the help of fuzzy logic

After determining the radius, the extent of the CH chance is calculated. In the fuzzy function, fuzzy rules use three fuzzy variables to calculate CH chance. They are:

Energy: the amount of remaining energy in the sensor node.

Neighbours: The number of neighbours who are located within the effect radius.

Probability: the likelihood that the sensor will lead the cluster.

Fuzzy rules used to calculate the sensor's chance of becoming cluster head are shown in Table 2:

Table 2

Fuzzy rules to calculate the sensor's chance for becoming cluster head.

NO.	Energy	Neighbor	Chance
1	Low	Little	Very Low
2	Low	Normal	Low
3	Low	Many	Rather Low
4	Medium	Little	Medium Low
5	Medium	Normal	Medium
6	Medium	Many	Medium High
7	High	Little	Rather High
8	High	Normal	High
9	High	Many	Very High

4.3. Clustering and determining CHs based on trust model

FCATM employs fuzzy logic in determining impact radius and calculating CH chance. This protocol elects CHs based on the greatest chance and the level of trust. Each sensor node in its radial range checks to see if other sensor nodes have any more chances, selects itself as cluster head, and publishes a notification message on the network. Other sensor nodes check, after receiving the message, if they are in the range of a few heads, to select the CH with the highest level of trust as their cluster head and send a joining message to this head.

4.4. Calculating trust level

In fuzzy logic, trust can be classified into two classes:

- Fuzzy direct trust
- Fuzzy indirect trust
-

4.4.1. Calculating fuzzy direct trust level

The trust extent of each sensor node is determined based on the history of communications between sensor nodes. In order to calculate the level of fuzzy direct trust or fuzzy membership function, each sensor node keeps track of successful and unsuccessful interactions with its neighbours. If, in the past, sensor A had successful interactions for S times and unsuccessful interactions for U times with sensor B, the fuzzy direct trust membership function for the relationship between A and B can be calculated as follows:

$$TD_{AB} = \frac{S_{AB}}{S_{AB} + U_{AB}} \tag{13}$$

4.4.2. Calculating fuzzy indirect trust level

In order to calculate the level of fuzzy indirect trust from sensor A to sensor B, each sensor node sends the calculated trust to other neighbours. When sensor A transmits the trust to sensor B from other neighbors, by using its trust level and the trust level received from other neighbors, it calculates the extent of trust to sensor B:

$$T_{AB} = w_1 \times TD_{AB} + w_2 \times \frac{\sum_{j \in N_B} TD_{jB}}{Length(N_B)} \tag{14}$$

Where NB represents the neighbors of sensor A that send their trust to sensor A. In addition, w1 and w2 are coefficients or weights of direct trust and neighbors' recommended trust, respectively, that have the following conditions:

$$w_1 + w_2 = 1 \quad (15)$$

The remarkable point is that at the beginning of the startup phase, the trust level to all sensors is considered equal and trust calculating is repeated at the beginning of each clustering step.

4.5. The pseudo code of FCATM

The pseudo code of the FCATM protocol is explained in Algorithm 1.

<u>Algorithm1. FCATM protocol</u>

Input: field dimensions, sensor position, sink position, maximum number of rounds, and initial energy

Output: A Clustered WSN, Average of Sensors' Energy, Average of Alive Nodes, Time of the First Sensor Death, Time of the First Cluster-Head Death

1. Computation of d0
2. Creation of the random sensor network
3. While all sensors have energy
4. Every time, the time of clustering
5. When the update finishes, do
6. 1. Clustering with consideration of the fuzzy-based radius
7. 2. Candidate Selected Cluster-Heads with Fuzzy-Based Chance
8. 3. Selection Cluster-Heads Based on Trust Model
9. Every time, beginning with the time
10. Calculate the average of the sensors' energy.
11. Calculate the average of alive nodes
12. Calculate the time of the first sensor death.
13. Calculate the time of the first cluster-head death.
14. End while

5. Simulation results and performance evaluations

We implemented FCATM, K-Means, FCM, subtractive clustering, and multi-objective fuzzy clustering algorithms in a MATLAB simulator to test and compare their performances in terms of average total remaining energy in the network, number of alive nodes, death time of the first sensor, and death time of the first CH. Simulation parameters are given in Table 3.

Table 3

Parameters	Values	Description
xm	300m	Field Dimensions: x maximum
ym	300m	Field Dimensions: y maximum
Sink.x	150m	Sink Position: x
Sink.y	400m	Sink Position: y
n	200	Number of Sensors
E0	0.5J	Initial Energy
Emp	0.0013pJ/bit/m4	Energy Consumed by power amplifier
Efs	10pJ/bit/m2	Energy Consumed by power amplifier
Eelec	50nJ/bit	Energy Consumed by transmitter and receiver circuits
ETX	50nJ/bit	Energy Consumed by radio electronics
ERX	50nJ/bit	Energy Consumed by radio electronics
EDA	5nJ/bit/signal	Energy Consumed for data aggregation
rmax	5000	maximum number of rounds

Simulation parameters.

Fig. 1 represents the location of elected CHs, and Fig. 2 represents clustering the nodes in the network with different colors in FCATM.

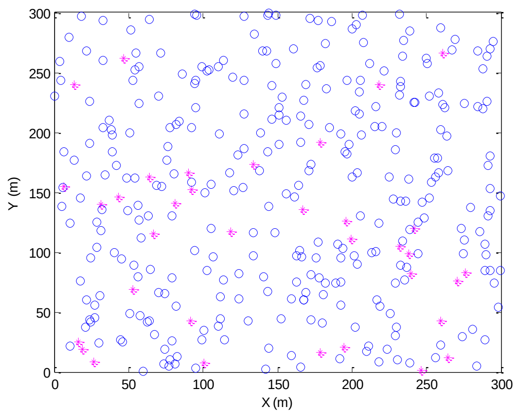


Fig 1. The location of elected CHs in FCATM.

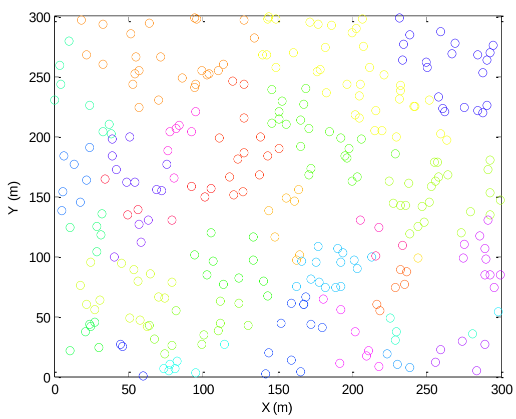


Fig 2. Clustering the nodes in the network with different colors in FCATM.

Fig. 3 compares the average total remaining energy in the algorithms after a certain period of 2500 rounds. It can be seen from the figure that the highest average of total remaining energy is obtained in FCATM, while the least is obtained in the K-Means algorithm for 200, 300, 400, 500, and 600 sensor nodes, respectively. Because in FCATM, determining the impact radius of the sensors based on fuzzy logic allows the best sensors to be selected as cluster heads, cluster heads away from the base station do not lose their energy quickly.

Fig. 4 compares the number of alive nodes after a certain period of 2500 rounds. Alive nodes are the nodes that have enough energy for data collection and processing. As can be seen from the figure, the number of alive nodes in FCATM is much higher than in the other algorithms because cluster head and cluster size election are based on fuzzy logic and load distribution is optimal, so the sensors will have a longer lifetime and the number of alive nodes is much higher than in the other algorithms.

Fig. 5 shows the comparison of the death time of the first sensor in the algorithms for 200, 300, 400, 500, and 600 sensor nodes, respectively. For each comparison, the death time of the first sensor in FCATM is the highest, while in the K-Means algorithm it is the least.

Figure 6 compares the initial CH's death time for several simulations by 200, 300, 400, 500, and 600 sensor nodes using various techniques. Because the cluster size in FCATM is inversely related to the CH's distance from the base station and distant CHs are the centres of smaller clusters, their energy consumption can be properly controlled, and the network loses them later, the death time of the first CH in FCATM is higher than that of the other algorithms for each comparison.

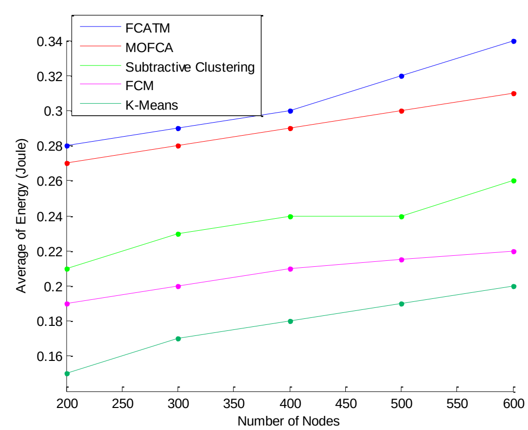


Fig 3. Comparison of the average total remaining energy with respect to the number of nodes.

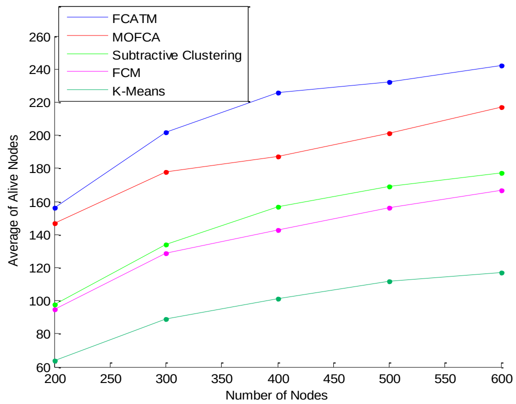


Fig 4. Comparison of the number of alive nodes after a certain period of 2500 rounds with respect to the number of nodes.

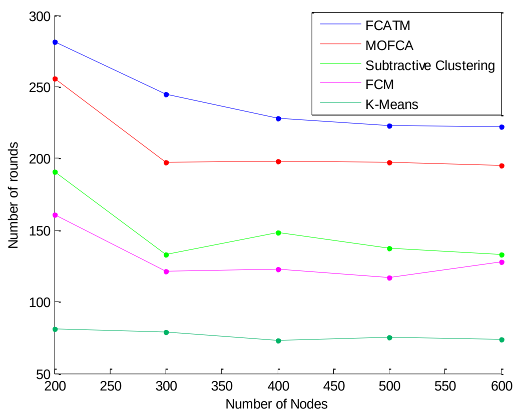


Fig 5. Comparison of the death time of the first sensor in the algorithms.

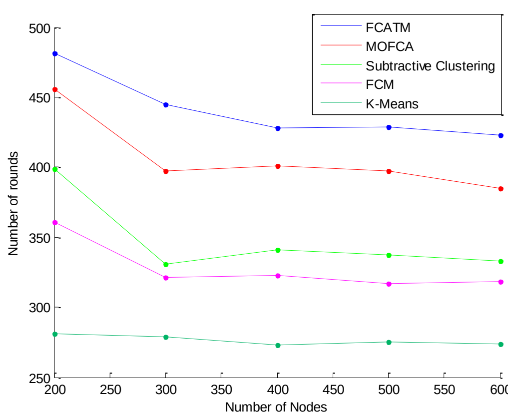


Fig 6. Comparison of the death time of the first CH in the algorithms.

6. Conclusions and future directions

In this study, a fuzzy clustering algorithm based on the trust model is proposed for wireless sensor networks. Our proposed algorithm, FCATM, considers nodes' energy levels, distance to the sink, impact radius, number of neighbours within the impact radius, and trust level parameters in selecting the cluster heads and joining the clusters while making use of fuzzy logic to overcome the uncertainties occurring in the WSN. We implemented FCATM in MATLAB software and compared the average of total remaining energy, the number of alive nodes, the death time of the first sensor, and the death time of the first CH with K-Means, FCM, subtractive clustering, and multi-objective fuzzy clustering algorithms for different simulation times and node numbers. We determined from simulation results that, FCATM performs best when compared with the other algorithms.
In future studies, the proposed study is planned to be organised for mobile wireless sensor networks. Also, applying useful parameters in the formation of clusters and using neural network topics in the algorithm will help to develop the proposed protocol. Moreover, using newer trust management systems to ensure security and determine the most reliable factors to evaluate the sensors' interactions can properly optimise the proposed protocol.

References

[1] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, F. Silva, "Directed diffusion for wireless sensor networking", ACM/ IEEE Transactions on Networking, 2002, vol. 11, pp. 2-16.

[2] Seyyit Alper Sert, Hakan Bagci, Adnan Yazici, "MOFCA: Multi-objective fuzzy clustering algorithm for wireless sensor networks", Applied Soft Computing, May 2015, vol. 30, pp. 151-165.

[3] A. Mahani, "Journal of Applied Research in Electrical Engineering A Novel Energy-Efficient Weighted Multi-Level Clustering Protocol Ebrahim Farahmand," vol. 1, no. 1, pp. 69–78, 2022, doi: 10.22055/jaree.2021.36169.1019

[4] N. I. Sarkar, D. P. Singh, and M. Ahmed, "A survey on energy harvesting wireless networks: Channel capacity, scheduling, and transmission power optimization," Electron., vol. 10, no. 19, pp. 1–20, 2021, doi: 10.3390/electronics10192342.

[5] M. Lotfinezhad, B. Liang, “Effect of partially correlated data on clustering in wireless sensor networks”, Proceedings of the IEEE International Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON), Citeseer, 2004, pp. 172–181.

[6] C. Nakas, D. Kandris, and G. Visvardis, “Energy efficient routing in wireless sensor networks: A comprehensive survey,” Algorithms, vol. 13, no. 3, pp. 1–65, 2020, doi: 10.3390/a13030072.

[7] Santosh.Irappa. Shirol, Ashok Kumar. N, Kalmesh.M. Waderhatti, “Advanced-LEACH Protocol of Wireless Sensor Network”, International Journal of Engineering Trends and Technology (IJETT), June 2013, Volume 4 Issue 6.

[8] O. Younis, S. Fahmy, “Distributed clustering in ad hoc sensor networks: a hybrid, energy-efficient approach”, Proceedings of the IEEE 23rd Joint Annual Conference of Computer and Communications Societies (INFOCOM), Hong Kong, vol.1, 2004; an extended version appeared in IEEE Transactions Mobile Computing, vol. 3, No. 4, 2004, pp. 366–379.

[9] L. Xuxun, “A survey on clustering routing protocols in wireless sensor networks”, Sens. J. 12 (8), 2012, pp. 11113–11153, http://dx.doi.org/10.3390/s120811113.

[10] Vaishali R. Patel, Rupa G. Mehta, “Impact of Outlier Removal and Normalization Approach in Modified k-Means Clustering Algorithm”, IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 5, No 2, September 2011.

[11] S. Suganthi, N. Umapathi, M. Mahdal, and M. Ramachandran, “Multi Swarm Optimization Based Clustering with Tabu Search in Wireless Sensor Network,” Sensors, vol. 22, no. 5, pp. 1–13, 2022, doi: 10.3390/s22051736.

[12] I. Gupta, D. Riordan, S. Sampalli, “Cluster-head election using fuzzy logic for wireless sensor networks”, Proceedings of the IEEE 3rd Annual Communication Networks and Services Research Conference, 2005, pp. 255–260.

[13] J. C. Bezdek, “Pattern Recognition with Fuzzy Objective Function Algorithms”, Plenum Press, 1981, New York.

[14] J. C. Bezdek and S. K. Pal, “Fuzzy Models for Pattern Recognition: Methods that Search for Structure in Data”, IEEE Press, 1992, New York.

[15] Nameirakpam Dhanachandra, Khumanthem Manglem and Yambem Jina Chuna, “Image Segmentation Using K-Means Clustering Algorithm and Subtractive Clustering Algorithm”, Procedia Computer Science 54, 2015, pp. 764-771.

[16] Wang Y, Zhao Q, Zheng D, “Energy-driven adaptive clustering data collection protocol in wireless sensor networks”, International conference on intelligent mechatronics and automation, Chengdu, China, 2004, p. 599–604.

[17] M. Majid et al., “Applications of Wireless Sensor Networks and Internet of Things Frameworks in the Industry Revolution 4.0: A Systematic Literature Review,” Sensors, vol. 22, no. 6, pp. 1–36, 2022, doi: 10.3390/s22062087.

[18] P. K. Mishra and S. Kumar, “Wireless Sensor Network for Underground Mining Services Applications,” Sens. Technol., pp. 452–478, 2020, doi: 10.4018/978-1-7998-2454-1

[19] S. Sadeghi, N. Soltanmohammadlou, and F. Nasirzadeh, “Applications of wireless sensor networks to improve occupational safety and health in underground mines,” J. Safety Res., vol. 83, pp. 8–25, Dec. 2022, doi: 10.1016/j.jsr.2022.07.016.

[20] F. Kyoomarsi, H. Khosravi, E. Eslami, P. K. Dehkordy, and A. Tajoddin, “Optimizing text summarization based on fuzzy logic,” Proc. - 7th IEEE/ACIS Int. Conf. Comput. Inf. Sci. IEEE/ACIS ICIS 2008, conjunction with 2nd IEEE/ACIS Int. Work. e-Activity, IEEE/ACIS IWEA 2008, pp. 347–352, 2008, doi: 10.1109/ICIS.2008.46.

Journal of Optimization of Soft Computing (JOSC)

Journal of Optimization of Soft Computing (JOSC)

Vol. 1, Issue 1, pp: (23-30), September-2023

Journal homepage: https://sanad.iau.ir/journal/josc

Social Spider Optimization Algorithm in Multimodal Medical Image Registration

Zahra Hossein-Nejad[1] and Mehdi Nasri [2*]

1. Department of Electrical Engineering, Sirjan brnach, Islamic Azad University Sirjan, Iran.

2. Department of Biomedical Engineering, Khomeinishahr branch, Islamic Azad University Isfahan, Iran.

Article Info

Article History:
Received 2023/8/13
Revised 2023/9/12
Accepted 2023/9/18

Keywords:
image registration, medical image processing, optimization, meta-heuristic algorithms, Social Spider Optimization.

**Corresponding Author's Email Address: nasri_me@iaukhsh.ac.ir*

Abstract

Image registration is one of the most essential applications of image processing, especially in medical. The purpose of image registration is to find the optimal parameters of conversion functions. The selection of appropriate optimization algorithms is very important in determining the optimal parameter. The Social Spider Optimization (SSO) algorithm is one of the meta-heuristic methods that prevents premature convergence. In this paper, medical image registration technique is suggested based on the SSO algorithm. The simulation results on Brain Web dataset affirm the suggested method outperforms classical registration methods in terms of convergence rate, execution time.

1. Introduction

Image registration is one of the most important branches of computer vision that is used in identifying changes, image fusion, image mosaics, image analysis and so on[1, 2]. Image registration is the process of finding geometric conversions images, taken at different imaging conditions [3]. Image registration approaches can be generally divided into two categories, feature-based approaches and intensity-based approaches[4, 5]. Intensity -based approaches versus feature-based approaches do not require preprocessing such as segmentation [6, 7].

Intensity-based methods usually include three steps: search space, similarity criteria, and search strategy [8, 9]. In the search space stage, based on the type of images and the distortion between the images, a suitable conversion for alignment is selected. Rigid and non-rigid conversion are examples of conversions. Similarity criterion is an index that measures the degree of similarity (correspondence) between two images (fixed, moving) that mutual information, correlation coefficients, joint entropy and etc are examples of similarity criteria. Mutual information (MI) criterion is one of the most popular similarity criteria, which is very effective in matching multimodal images [10, 11]. The search strategy step is one of the most important steps in intensity-based methods because the performance of this step directly affects the matching results [12, 13].In this step, the parameters of the optimal conversion function are computed using the optimization algorithm for image alignment. Optimization algorithms can be broadly divided into local and global categories [9]. Local algorithms have high

exploration power and these algorithms have less execution time. Local optimizers may find local optimizers instead of global optimizers, which is one of the disadvantages of these optimizers.
In [14], the simulated annealing algorithm is used for the noise image registration. In [15], genetic algorithm and mutual information similarity criterion are used to register the multimodal images of the brain. In [16], Genetic Algorithm (GA) and Affine conversion have used for image registration. In [6], the Particle Swarm Optimization (PSO) algorithm and the mutual information criterion have used in the medical multimodal image registration. In [17], the genetic algorithm and the simulated annealing algorithm are compared to retina image registration, which shows the genetic algorithm has better performance in terms of convergence rate and speed. In [18], grey-wolf-based Wang's demons algorithm is used for retinal images registration. In [7], the Bat algorithm and Grey Wolf optimization with the criterion of similarity of mutual information is used to register medical images such as brain and retina. In [19], the firefly algorithm with the correlation coefficient similarity metric is used for image registration. This algorithm has better accuracy and speed in image registration than PSO algorithm. In [20], the Accelerated Particle Swarm Optimization (APSO) algorithm with the criterion of similarity of mutual information for image registration is presented. This algorithm has a better speed but less accuracy in image registration than the firefly algorithm. In [21], artificial bee colony algorithm with mutual information similarity metric is proposed for image registration. This method is more accurate than other methods. In [22], PSO and the Artificial Bee Colony algorithm (ABC) are used in medical images registration. The results show that ABC algorithm is more accurate than PSO algorithm, but this algorithm has more execution time. In [8], ant colony algorithm and mutual information are presented in registration MRI and CT images of the brain. In [10], ant colony algorithm and Differential total variation (DTV) are proposed for search space and similarity metric in multimodal remote sensing images registration, respectively. In [23] , the PSO method and the sequential quadratic programming (SQP) method are combined in the search space, called the PSOSQP method. This method has better accuracy and speed in image registration than other methods such as GA, PSO and ABC. In [11], ant colony algorithms have been improved for image registration, which is of better quality and speed than other methods such as classic ant colony algorithm, PSO.
Choosing a suitable method for search strategy is very important because the speed of convergence, execution time and finally reaching the optimal answer depends on the type of search strategy. Various meta- heuristic algorithms have been used so far for search strategy in image registration. The social spider optimization algorithm is another meta-heuristic algorithm. This algorithm was presented by Erik Cuevas et al based on the cooperative characteristics of the social spider [24]. The social spider algorithm has been widely used in various applications such as image contrast enhancement [25], anti-islanding protection [26], text psychology analysis [27], and energy theft detection[28] .This algorithm unlike other algorithms as PSO [29], GA [30], Cuckoo Search (CS) [31], ABC [32], Harmony Search (HS) [33], and Social Network Optimization (SNO) [9], prevents premature convergence to local optimum solutions or a limited balance between exploration and exploitation [34]. These advantages motivate the use of the SSO for medical images registration. In this article, the SSO algorithm is proposed for finding the optimal parameters of affine conversion in medical image registration. Another innovation is the study of similarity criteria such as mutual information, normalization of mutual information (NMI), and Sum of Squared Differences (SSD) in the registration of brain images.
Organization of the remainder of the paper is as follows. In Section II, the suggested algorithms are presented and in Section III, the tests results are studied. Section IV and Section V, are discussion and conclusions are studied, respectively.

III- METHOD

The image registration consists to search for the optimal geometric conversion between moving image (s) and fixed image (r) according to (1).

$$T^* = \arg\max S(r, T_a(s)) \tag{1}$$

In (1), T_a is the transformation matrix by transformation parameters a, and S is fitness is maximized when the r image is completely aligned with $T_a(s)$ image.To register the transformed moving image $T_a(s)$ to the fixed image r, the set of parameters a that maximizes the fitness function S needs to be estimated by search strategy .The conversion function, similarity metric, and search strategy used are described below.

A. Affine Conversion Function

An appropriate conversion type should be selected for image registration according to the type of images and the deviation between the images. Conversions are broadly classified into two categories, rigid conversion and non-rigid conversion [7, 20]. The Affine conversion is one of the non-rigid transformations that includes four parameters, scaling (S), translation (T), rotation (R) and shear (SH) according to.(2).

$$AC = S \times T \times R \times SH$$

$$AC = \begin{bmatrix} S_1 & 0 & 0 \\ 0 & S_2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & sh_1 & 0 \\ sh_2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

B. Similarity Metric

The similarity metric measures the similarity between the fixed image and the moving image, which is used as a fitness function in image registration algorithms. Similarity metrics are mutual information, correlation, entropy, etc. The choice of appropriate similarity metric depends on the type of image [35]. In this paper, fitness functions such as mutual information according to (3) and normalized mutual information according to (4) and SSD according to (5) are used.

$$MI(A,B) = H(A) + H(B) - H(A,B) \quad (3)$$

$$NMI(A,B) = \frac{H(A)+H(B)}{H(A,B)} \quad (4)$$

$$SSD = \sum_{i=1}^{n} \left(A(z_i) - B(Tz_i)\right)^2 \quad (5)$$

In (3-5), $H(A,B)$ is the joint entropy, $H(A)$ and $H(B)$ are marginal entropy in the fixed image and marginal entropy in the moving image, respectively.

C. search strategy by SSO Algorithm

The social spider optimization algorithm was introduced in 2013 [24]. Social spider-seeking behavior can be described as the collective movement of spiders toward the food source. Every spider on the web has a position and fitness, which indicates the potential to find a food source in that position. The spider can move freely on the web but it cannot leave the web. When a spider moves to a new position, it produces a vibration that propagates across the web. The social members of the spider are divided into males and females, in which the number of females is more than males. The number of female spiders makes up about 90% of the population. The spider interacts with the vibration of the strings on the web. The vibration that a female spider receives is according to (6), depending on the size of the spider and the distance of the spiders, regardless of the type of spider.

$$Vib_{i,j} = w_j \cdot e^{-d_{i,j}^2} \quad (6)$$

In (6), w_j is the weight of spider j and $d_{i,j}$ is the Euclidian distance between spiders i and j.
Female spiders are generally able to sense three vibrations from spiders: 1- The closest spider with a higher weight (v_{ibci}), 2- The earliest spider in the community (v_{ibbi}),3-the closest male spider to the female spider (v_{ibfi}).
The female spider may have a gravitational or repulsive motion toward the source of the vibration. If the spider is attracted to the vibration, its position is updated according to (7), and if the spider moves away from the vibration position, its position is updated according to (8).

$$f_i^{k+1} = f_i^k + \alpha \cdot Vibc_i \cdot \left(s_c - f_i^k\right) + \beta \cdot Vibb_i \cdot \left(s_b - f_i^k\right) + \delta \cdot \left(rand - \frac{1}{2}\right) \quad (7)$$

$$f_i^{k+1} = f_i^k - \alpha \cdot Vibc_i \cdot \left(s_c - f_i^k\right) - \beta \cdot Vibb_i \cdot \left(s_b - f_i^k\right) + \delta \cdot \left(rand - \frac{1}{2}\right) \quad (8)$$

In (7-8), $\alpha, \beta, \delta,$ and $rand$ are random numbers between $[0,1]$ and k shows the iteration number. s_c, s_b are the closest members to the i, who have the most weight and the best member in the whole population S, respectively.

Male spiders are divided into two groups, dominant spiders and recessive spiders, according to the biological behavior of social spiders. Dominant male spiders are heavier and female spiders are more likely to be attracted to its vibrations, but recessive male spiders tend to move toward the center of the population and use food resources wasted by the dominant spider as a strategy. The position of the dominant spider is updated according to (9) and the position of the defeated spider is updated according to (10).

$$m_i^{k+1} = m_i^k + \alpha \cdot V_{ibfi} \cdot \left(s_f - m_i^k\right) + \delta \cdot \left(rand - \frac{1}{2}\right) \quad (9)$$

$$m_i^{k+1} = m_i^k + \alpha \cdot \left(\frac{\sum_{h=1}^{N} m_h^k \cdot w_{Nf+h}}{\sum_{h=1}^{N} w_{Nf} + h} - m_i^k \right) \quad (10)$$

In (9-10), $\alpha, \beta,$ and $rand$ are random numbers between $[0,1]$ and s_f is the nearest female to male i whereas $\left(\frac{\sum_{h=1}^{N} m_h^k \cdot w_{Nf+h}}{\sum_{h=1}^{N} w_{Nf+h}} \right)$ is the weighted mean of the male population M.

II. IMPLEMENTATION AND EXAMINATION OF RESULTS

To evaluate the function of the suggested approach, three sets of tests are performed with the classical algorithms such as GA, PSO. The database used in this article contains images with different modalities such as a set of T1, T2 and PD-weighted MR images of size $256 \times 256 \times 32$ voxels from the BrainWeb database [36]. In the first set of experiments, the performance of the suggested approach on mono-modal images of the brain is investigated. In the second test set, the performance of the suggested approach on multi-modal brain of images is examined. In the third test set, effect of different Fitness functions in image registration is studied. The results of the tests are checked by convergence rate, execution time, and RMSE and average MI.

A. Registration performance of the suggested approach on mono-modal images

In the first test, the proposed algorithm is applied on mono-modal brain images, and its function is investigated (Fig.1 and Table1) .

TABLE I- Results of GA-SSD, PSO-SSD and SSO-SSD in mono-modal image registration

Type image	Algorithm	RMSE
T1-T1	GA-SSD	5.481
	PSO-SSD	4.965
	SSO-SSD	**3.387**
T2-T2	GA-SSD	6.142
	PSO-SSD	4.854
	SSO-SSD	**4.307**

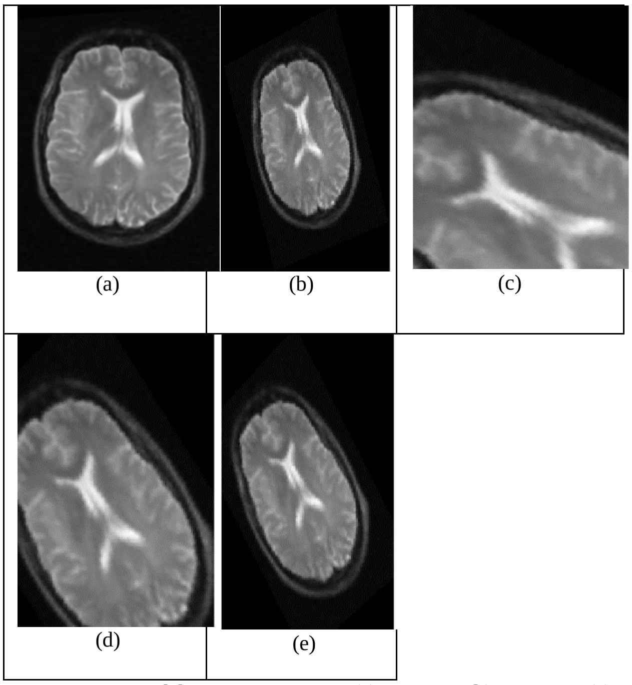
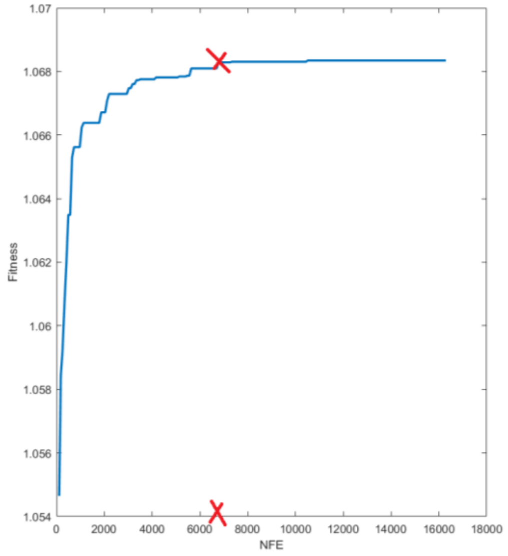


Fig.1. mono- modal image registration, (a)T2 image, (b) T2 image, (c) mono-modal image registration by GA, (d)mono-modal image registration by PSO (e) mono- modal image registration by SSO

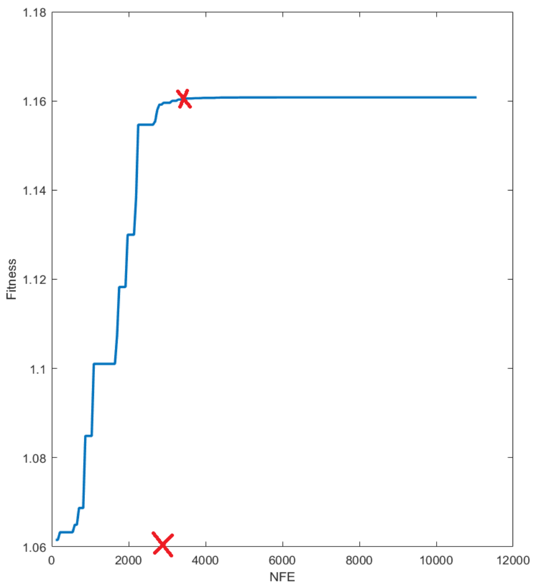
According to Fig.1, it can be concluded that the SSO has a better function in image registration than other meta-heuristic algorithms.

B. Registration performance of the suggested approach on multimodal images

In this test, the multimodal brain images were used to assess the functionality of the proposed method in registration process; the results are shown in Fig. 2 &3 and Table 2.



(a)



(b)

Fig.2. Graph of number functions evaluated in terms of fitness functions, (a) GA - MI, (b) SSO algorithm-MI

Fig.2 is a graph of the number of functions evaluated (NFE) in terms of the mutual information fitness function. To compare these two algorithms (GA, SSO), the same population size and number of iterations are considered. According to the graphs, it was found that the SSO algorithm has reached a better optimization value than the GA. On the other hand, the SSO algorithm is far more powerful than the GA in terms of convergence speed by comparing the NFE. For example, to optimize this problem, the number of evaluated functions for the GA is 7000 and for the social spider algorithm is 3500, which indicates that the GA results in image registration solving fewer equations than the GA. This demonstrates the effective efficiency of the SSO algorithm in brain images registration.

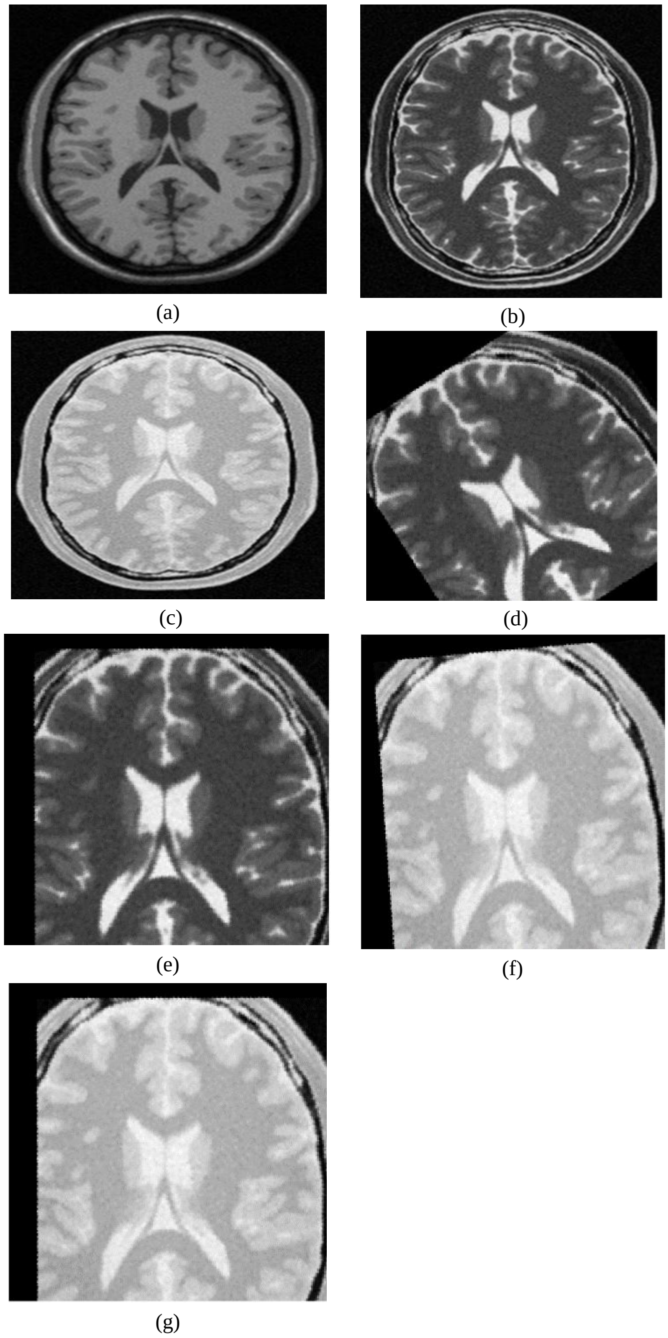


Fig.3. image registration, (a)T1 image, (b)T2image, (c) PD image (d) image (T1&T2) registration by GA, (e) image (T1&T2) registration by SSO, (f) image(T2&PD) registration by GA (g) image(T2&PD) registration by SSO

The unsuccessful registration of the GA is due to not finding the optimal cconversion parameters (Fig.3, d). In the proposed methods, due to finding the most optimal conversion parameters, the registration process is better done (Fig.3 e.g.).

TABLE II-MAE comparison between genetic-MI, PSO-MI, and SSO-MI

Image	GA -MI				PSO-MI				Proposed SSO-MI			
	t_x	t_y	θ	times	t_x	t_y	θ	times	t_x	t_y	θ	times
T1-T2	0.303	0.247	0.0026	418.33	0.301	0.228	0.0023	341.17	0.272	0.225	0.0019	294.44
T2-PD	0.383	0.296	0.0025	513.28	0.306	0.224	0.0022	427.20	0.295	0.217	0.002	350.55
Average	0.343	0.271	0.0025	465.80	0.3035	0.226	0.0024	384.19	**0.2835**	**0.221**	**0.00195**	**322.50**

C. The effect of different Fitness functions in image registration

In this test, five pairs of multimodal images and five pairs of mono modal images were used to evaluate different fitness function (Table 2).

Table II-Evaluate image registration based on SSO algorithm using different fitness functions

Type image	Algorithm	Average-MI
Mono modal	SSO-MI	1. 120
	SSO-NMI	1.191
	SSO- SSD	**1.138**
Multi modal	SSO-MI	**1.16**
	SSO-NMI	1.141
	SSO-SSD	1.119

III. **Discussion**

In this part, registration results of the suggested method and GA, and PSO algorithms in the brain database for all evaluation criteria will be reviewed.

The image registration performance according to the RMSE from best to worst is the SSO, PSO, and the genetic algorithm, respectively (Table I). Table II shows the median absolute error (MAE) value for each Affine conversion parameter in each algorithm that parameter t_x , the order of algorithm performance from best to worst is SSO (0.2835), PSO (0.3035) and GA (0.343). For parameter t_y , and θ the order of algorithm from best to worst is SSO (0.221, 0.00195), PSO (0.226, 0.0024) and GA(0.271, 0.0025), respectively. We can conclude that the SSO performs best for the estimation of all parameters. The SSO algorithm has the highest speed and the GA has the lowest speed in medical images registration (Table II). In Table III, different fitness functions for mono-modal and multimodal images registration were investigated. The mutual -information fitness function for multimodal image registration and the SSD fitness function for mono-modal images registration performed better.

IV. **Conclusion**

In this article, SSO-based image registration is proposed using similarity metrics such as MI, NMI, and SSD. First, the performance of these similarity metrics is compared with each other, and then the performance of the SSO is compared with other optimization algorithms such as GA and PSO in brain image registration. The results show that the metric similarity of mutual information has a better performance in multimodal images registration and the function of SSO is better in terms of speed and quality of image registration than GA and PSO algorithm. We will try to use this suggested approach in other image registration such as remote-sensing image and natural image in future works.

--

references

[1] Z. Hossein-Nejad and M. Nasri, "Adaptive RANSAC and extended region-growing algorithm for object recognition over remote-sensing images," Multimedia Tools and Applications, 2022.

[2] Z. Hossein-Nejad and M. Nasri, "Clustered redundant keypoint elimination method for image mosaicing using a new Gaussian-weighted blending algorithm," The Visual Computer, 2021/07/19 2021.

[3] Z. hossein-nejad and M. Nasri, "Clustered Redundant Keypoint Elimination SURF method in MRI Image Registration based on Alpha-Trimmed Relationship," Frontiers in Biomedical Technologies, 2022.

[4] Z. Hossein-Nejad and M. Nasri, "Image Registration Based on Redundant Keypoint Elimination SARSIFT Algorithm and MROGH Descriptor," in 2022 International Conference on Machine Vision and Image Processing (MVIP), 2022, pp. 1-5.

[5] Z. Hossein-Nejad and M. Nasri, "Retinal Image Registration based on Auto-Adaptive SIFT and Redundant Keypoint Elimination Method," in

2019 27th Iranian Conference on Electrical Engineering (ICEE), 2019, pp. 1294-1297.
[6] L. Rundo, A. Tangherloni, C. Militello, M. C. Gilardi, and G. Mauri, "Multimodal medical image registration using particle swarm optimization: a review," in 2016 IEEE Symposium Series on Computational Intelligence (SSCI), 2016, pp. 1-8.
[7] K. Shaw, P. Pandey, S. Das, D. Ghosh, P. Malakar, and S. Dhabal, "Image Registration using Bio-inspired Algorithms," in 2020 IEEE 1st International Conference for Convergence in Engineering (ICCE), 2020, pp. 330-334.
[8] H. Rezaei, M. Shakeri, S. Azadi, and K. Jaferzade, "Multimodality image registration utilizing ant colony algorithm," in 2009 Second International Conference on Machine Vision, 2009, pp. 49-53.
[9] M. P. Wachowiak, R. Smolíková, Y. Zheng, J. M. Zurada, and A. S. Elmaghraby, "An approach to multimodal biomedical image registration utilizing particle swarm optimization," IEEE Transactions on evolutionary computation, vol. 8, pp. 289-301, 2004.
[10] Y. Wu, W. Ma, Q. Miao, and S. Wang, "Multimodal continuous ant colony optimization for multisensor remote sensing image registration with local search," Swarm and Evolutionary Computation, vol. 47, pp. 89-95, 2019.
[11] Y. Tian and H. D. Ma, "Image registration based on improved ant colony algorithm," in Advanced Materials Research, 2013, pp. 683-686.
[12] Y. Zhuang, K. Gao, X. Miu, L. Han, and X. Gong, "Infrared and visual image registration based on mutual information with a combined particle swarm optimization–Powell search algorithm," Optik, vol. 127, pp. 188-191, 2016.
[13] H. Ghorbani and A. A. Beheshti, "Multiresolution registration of multitemporal remote sensing images by optimization of mutual information using a simulated annealing based Marquardt-Levenberg technique," in 2007 International Conference on Intelligent and Advanced Systems, 2007, pp. 685-690.
[14] I. Maglogiannis and E. Zafiropoulos, "Automated medical image registration using the simulated annealing algorithm," in Hellenic Conference on Artificial Intelligence, 2004, pp. 456-465.
[15] H. Zhang, X. Zhou, J. Sun, and J. Zhang, "A novel medical image registration method based on mutual information and genetic algorithm," in International Conference on Computer Graphics, Imaging and Visualization (CGIV'05), 2005, pp. 221-226.
[16] R. Suganya and S. Rajaram, "Mutual information and genetic algorithm based registration of MRI brain images," Canadian Journal on Biomedical Engineering and Technology, vol. 2, pp. 14-17, 2011.
[17] G. K. Matsopoulos, N. A. Mouravliansky, K. K. Delibasis, and K. S. Nikita, "Automatic retinal image registration scheme using global optimization techniques," IEEE Transactions on Information Technology in Biomedicine, vol. 3, pp. 47-60, 1999.
[18] S. Chakraborty, R. Pradhan, A. S Ashour, L. Moraru, and N. Dey, "Grey-Wolf-Based Wang's Demons for retinal image registration," Entropy, vol. 22, p. 659, 2020.
[19] S. Chakraborty, N. Dey, S. Samanta, A. S. Ashour, and V. E. Balas, "Firefly algorithm for optimized nonrigid demons registration," in Bio-inspired computation and applications in image processing, ed: Elsevier, 2016, pp. 221-237.
[20] C.-L. Cocianu and A. Stan, "New evolutionary-based techniques for image registration," Applied Sciences, vol. 9, p. 176, 2019.
[21] Y. Qin, H. Hu, Y. Shi, Y. Liu, and H. Gao, "An artificial bee colony algorithm hybrid with differential evolution for multi-temporal image registration," in 2016 35th Chinese Control Conference (CCC), 2016, pp. 2734-2739.
[22] D. Sarvamangala and R. V. Kulkarni, "Swarm intelligence algorithms for medical image registration: A comparative study," in International Conference on Computational Intelligence, Communications, and Business Analytics, 2017, pp. 451-465.
[23] Y. Zhang and L. Wu, "Rigid image registration by PSOSQP algorithm," Advances in Digital Multimedia, vol. 1, pp. 4-8, 2012.
[24] E. Cuevas, M. Cienfuegos, D. Zaldívar, and M. Pérez-Cisneros, "A swarm optimization algorithm inspired in the behavior of the social-spider," Expert Systems with Applications, vol. 40, pp. 6374-6384, 2013.
[25] L. Maurya, P. K. Mahapatra, and A. Kumar, "A social spider optimized image fusion approach for contrast enhancement and brightness preservation," Applied Soft Computing, vol. 52, pp. 575-592, 2017.
[26] R. Vijay and V. Priya, "Anti-islanding protection of distributed generation based on social spider optimization technique," International Journal of Advanced Engineering Research and Science, vol. 4, p. 237193, 2017.
[27] A. Hosseinalipour, F. S. Gharehchopogh, M. Masdari, and A. Khademi, "Toward text psychology analysis using social spider optimization algorithm," Concurrency and Computation: Practice and Experience, vol. 33, p. e6325, 2021.
[28] D. R. Pereira, M. A. Pazoti, L. A. Pereira, D. Rodrigues, C. O. Ramos, A. N. Souza, et al., "Social-Spider Optimization-based Support Vector Machines applied for energy theft detection," Computers & Electrical Engineering, vol. 49, pp. 25-38, 2016.

[29] J. Kennedy and R. Eberhart, "Particle swarm optimization," in Proceedings of ICNN'95-international conference on neural networks, 1995, pp. 1942-1948.

[30] M. Kumar, M. Husain, N. Upreti, and D. Gupta, "Genetic algorithm: Review and application," Available at SSRN 3529843, 2010.

[31] A. H. Gandomi, X.-S. Yang, and A. H. Alavi, "Cuckoo search algorithm: a metaheuristic approach to solve structural optimization problems," Engineering with computers, vol. 29, pp. 17-35, 2013.

[32] D. Karaboga and B. Basturk, "A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm," Journal of global optimization, vol. 39, pp. 459-471, 2007.

[33] X.-S. Yang, "Harmony search as a metaheuristic algorithm," in Music-inspired harmony search algorithm, ed: Springer, 2009, pp. 1-14.

[34] A. Luque-Chang, E. Cuevas, F. Fausto, D. Zaldivar, and M. Pérez, "Social spider optimization algorithm: modifications, applications, and perspectives," Mathematical Problems in Engineering, vol. 2018, 2018.

[35] R. Panda, S. Agrawal, M. Sahoo, and R. Nayak, "A novel evolutionary rigid body docking algorithm for medical image registration," Swarm and Evolutionary Computation, vol. 33, pp. 108-118, 2017.

[36] R.-S. Kwan, A. C. Evans, and G. B. Pike, "MRI simulation-based evaluation of image-processing and classification methods," IEEE transactions on medical imaging, vol. 18, pp. 1085-1097, 1999.

Journal of Optimization of Soft Computing (JOSC)

Journal of Optimization of Soft Computing (JOSC)

Vol. 1, Issue 1, pp: (31-38), September-2023

Journal homepage: https://sanad.iau.ir/journal/josc

An Enhanced Rotor Side Converter Control of DFIG-Based Wind Turbines For Improving LVRT Capability During Balanced Grid Faults

Hamid Rahimi Esfahani

Department of Electrical Engineering, Lenjan Branch, Islamic Azad University, Isfahan, Iran

Article Info

Article History:
Received 2023/10/7
Revised 2023/10/24
Accepted 2023/10/30

Keywords:
Demagnetization control method, DFIG, LVRT, Balanced three-phase fault

**Corresponding Author's Email Address: Hamid.Rahimi@iau.ac.ir*

Abstract

Among variable speed wind turbines the use of doubly-fed induction generator based wind turbines are very common. However, due to their stator direct connection to the grid, they are very sensitive to grid faults. This paper analyzes the low voltage ride-through (LVRT) capability enhancement under the three-phase balanced grid fault. Therefore the variations of 5 parameters during balanced grid fault are Analyzed. These parameters are stator voltage, rotor current, stator current, rotor speed, and the DC link voltage. To improve the LVRT capability an enhanced demagnetization control method is proposed. This method shortens the natural stator current time constant and approximately immunes the system against parameter variations as shown in simulation results. To show the effectiveness of the proposed method, the results are compared with one of the best previous demagnetization control methods.

1. Introduction

Today, wind energy plays an important role in modern renewable energy sources. The advantages of doubly-fed induction generator-based wind turbines (DFIG WTs), such as variable-speed -constant-frequency based operation and decoupled control of active and reactive power has made them one of the most employed WT types at the range of 1 MW and more. Furthermore, the rotor power is managed only by its converter which can be rated at about one-third of the whole generator rating [1-3], which brings the advantages of being light, low cost with small losses compared to WTs with a full-scale converter. However, the Direct DFIG Stator connection to the grid is its main disadvantage, which makes it very sensitive to any voltage disturbances such as low voltage and fault. At the beginning of wind energy employment, the generator unit was disconnected from the grid as soon as occurring a fault. By increasing the penetration of wind turbine systems, during occurring a fault it is required that wind farms to remain connected to the grid for certain time duration and under special criteria (to help the power system stability by supplying reactive power), which is called Low Voltage Ride Through (LVRT) requirement [4].

To prevent the disconnection of wind farms during network disturbances, The grid code requirements for LVRT are imposed by many countries. To define the requirements on the LVRT capability of grid-connected wind farms, China Electric Power Research Institute (CEPRI) developed revising technical rule for connecting wind farms to the power system (GB/T 19963-2005) via inserting specific necessities on the LVRT curve. This rule has been applied since 30 June 2012 [5]. Fig. 1 shows the LVRT requirements for wind farms. According to this new standard, WTs should be able to preserve operating for 625 mS when the voltage of PCC falls to 20% of the nominal voltage. Also in the new standard, WTs should stay connected to the power grid and keep on operating when 90% of the nominal PCC voltage is retrieved in 2 seconds [5].

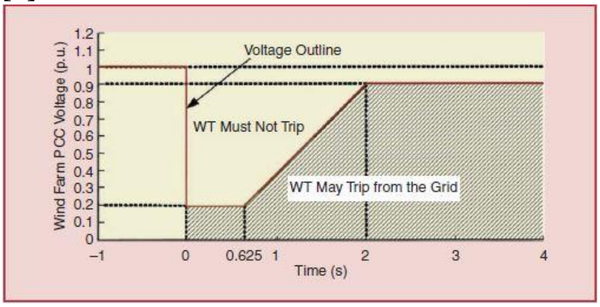


Fig. 1. LVRT requirements for wind farms [5]

Among different proposed methods for improving LVRT capability, the employment of a crowbar [6] is very common. On detecting of the rotor overcurrent, the crowbar short circuits the rotor terminal and the trigging signal of the Rotor Side Converter (RSC) is blocked simultaneously, to be protected. But, it changes the DFIG right into an induction generator which absorbs reactive power, also deteriorating the grid fault. As a result, dynamic VAR compensators are occasionally installed at the DFIG terminals to provide reactive power for the duration of grid faults [7]. Instead of using the crowbar, other methods based on extra power electronic equipment are proposed, such as suggesting the schemes with an additional grid side converter (GSC) [8], [9], dynamic voltage restorer [10], [11] and energy storage system [12]. However, the high cost and control complexity discourages their utilization. The implementation of a stator side passive impedance network [13], rotor side dynamic resistor [14] and rotor side reactor [15] are also proposed. On this manner, the wind turbine can stay connected to the grid during occurring voltage dips. But, sizing of the impedance is made either experimentally or via simulation, and therefore it is quite difficult to optimize the whole system.
In some references, due to easy implementation, solutions based on advanced control methods are suggested. It's been confirmed that Feed-Forward Transient Current Control (FFTCC) can restrict transient rotor current during the low voltages and faults, Resulting in a minimum incidence of crowbar interruptions [16], [17]. But, the torque is constant to zero and the reactive power needs to be drawn from the grid. to counteract the natural and negative sequence components of the stator flux, demagnetization control is proposed in [18]. And it's been verified that the control should promote natural stator flux damping [19]. An aggregate of demagnetization control and active crowbar is proposed to make the crowbar activation time shorter [20]. But, the demagnetization control in the aforementioned researches is sensitive to system parameter variation since stator resistance information is wanted, in spite that the sensitivity can be decreased when demagnetization control is blended with virtual resistance [21]. When a fault occurs, natural stator flux is produced and in an unbalanced fault, extra negative sequence stator flux is added. The two aforementioned fluxes induce a very high electromotive force (EMF) in the rotor circuit leading to a large transient rotor current and finally the destruction of the converter and disconnection of the system.
This paper deals about balanced grid faults and its effects on RSC converter. Therefore, an improved demagnetising control method is proposed to compensate for the harmful effects of the high EMF and large currents in the rotor circuit. In this method, during the fault, the variations of the five parameters of stator terminal voltage (V_t), stator current (I_s), rotor current (I_r), rotor speed (ω_r) and dc-link voltage (V_{dc}), are analysed to see the effectiveness of the proposed method for improving the LVRT capability. Finally, the proposed method is compared to the proposed method in [22] and the effect of changing machine parameters on the LVRT capability in the proposed method is analysed at the end.
This paper is organized as follows. Section 2 analyzes system behaviour during a balanced grid fault. It shows that the natural EMF produced during the fault is influenced by fault duration and stator flux time constant. Section 3 introduces the proposed demagnetization control method, and the influence on the LVRT performance is discussed in Section 4 through simulation results..It is proved that system LVRT capability can be improved by reducing the time constant of the natural stator current during a balanced grid fault. The conclusion is presented in section 5.

2. The Behaviour of the DFIG Under Symmetrical Grid Faults

Concerning the stator reference frame, the stator voltage space vector v_s is expressed as

$$v_s = r_s i_s + \frac{d}{dt}\psi_s \tag{1}$$

Where v_s, r_s, i_s and ψ_s represent the stator voltage, resistance, current and flux space vector, respectively and $\frac{d}{dt}$ is the derivative operator. In the case of symmetrical three-phase grid faults, the stator voltage vector can be expressed as

$$v_s = V_p e^{j\omega_s t} \tag{2}$$

where V_p is the amplitude of positive sequence components and ω_s is the synchronous speed. If r_s is neglected, according to (1) and (2), the steady-state component of the stator flux during grid faults can be expressed as

$$\psi_{ss} = \frac{V_p e^{j\omega_s t}}{j\omega_s} \quad (3)$$

where ψ_{ss} is the steady-state component of the stator flux. Due to the continuous variation of the flux from its initial state to its steady-state, the total stator flux includes the transient component and steady-state component during grid faults is expressed as

$$\psi_s = \psi_{ss} + \psi_{st} = \frac{V_p e^{j\omega_s t}}{j\omega_s} + \Psi_{st} e^{-\frac{t}{\tau}} \quad (4)$$

where ψ_{st} and Ψ_{st} are the transient component of the stator flux and its initial value, respectively, and τ is the time constant.
Moreover,

$$\psi_s^r = \psi_s e^{-j\omega_r t} \quad (5)$$

where superscript "r" denotes that the equations are expressed in the rotor reference frame and ω_r is the angular velocity of the rotor.
The rotor circuit EMF, induced by a stator flux, can be expressed as

$$E = \frac{L_m}{L_s} \cdot \frac{d}{dt} \psi_s^r \quad (6)$$

where L_s and L_m are the stator self-inductance and the mutual inductance between the stator and rotor, respectively.
Based on (4), (5) and (6), the EMF E during the grid fault can be derived as

$$E = \frac{L_m}{L_s} s V_p e^{js\omega_s t} + \frac{L_m}{L_s} (\frac{1}{\tau} + j\omega_r) \Psi_{st} e^{-\frac{t}{\tau}} e^{j\omega_r t} \quad (7)$$

where s is the slip rate. By neglecting $\frac{1}{\tau}$ [18], (7) is simplified to (8).

$$E = \frac{L_m}{L_s} s V_p e^{js\omega_s t} + \frac{L_m}{L_s} (1 - s) j\omega_s \Psi_{st} e^{-\frac{t}{\tau}} e^{j\omega_r t} \quad (8)$$

According to equation (8), E consists of two parts. the first part and the second part, are induced by positive and dc components of the stator flux, respectively. Because of small s, the second term induced by the DC component of the stator flux during the faults can cause overvoltages/overcurrents at the rotor side. under symmetric fault, the second term of the stator flux consists of exponentially decayed dc component is added compared to the pre-fault state. The decayed dc component is static concerning to stator windings and rotates at rotor speed concerning to rotor windings; according to the second term in (8), the EMF induced by the dc flux component is proportional to (1−s): this EMF is much higher than the EMF corresponding to the normal operation which is proportional to s; so the total EMF is much higher than that in normal operation and it may exceed the maximum voltage of the rotor converter. The induced large voltages in the rotor windings result in large currents into the rotor circuit, which causes severe damage to RSC and the large increase of the dc-link voltage. Thus, the proposed method in this paper is intended to limit rotor overvoltage and consequently rotor current or decrease the instability rate during the fault. In the next part, the proposed method is introduced.

3. Proposed LVRT Capability Enhancement Method

Though the vector control is commonly used to control the DFIG [23-25], the control method in this paper is focused only on the RSC and nor the GSC. In the following subsection, the proposed demagnetizing method is described. The advantage of this method compared to previous demagnetizing control methods is its lower dependence on machine parameter changes. To show this independence, the comperative simulation results are presented in section 4.

3.1. Control Method Based On Enhanced Demagnetizing Current

The traditional method according to [18-20], was based on mitigating the influence of natural stator flux through decreasing the amplitude of natural stator flux. The main drawback of this method is its sensitivity to the system parameter variation. To lessen this sensitivity, in this paper an improved demagnetization control method is suggested which does not need the system parameter information. In this method, instead of decreasing the natural stator flux amplitude, the time constant of natural stator current is decreased to cause the faster decay of the natural stator flux. Note that, the stator current is proportional to the stator flux, and any changes in the current cause the same effect on the flux. According to equations (9-11), the relation between the natural stator current and the natural stator flux is obtained.
Based on the DFIG model the stator flux is described as

$$\psi_s = L_s i_s + L_m i_r \quad (9)$$

Where, i_r represents the rotor current. The layout of the control approach is shown in Fig. 2. Demagnetizing rotor current, which opposes the natural stator current, is injected into the rotor circuit during the occurence of a balanced three-phase grid fault.

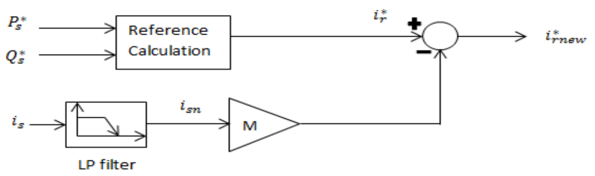


Fig.2. The block diagram of the proposed demagnetization control.

The reference is given as

$$i_{rn} = -M i_{sn} \quad (10)$$

where M is a positive demagnetization coefficient and i_{rn}, i_{sn}, are natural rotor current and natural stator current respectively. By combining (9) and (10), equation (11) is written as

$$i_{sn} = \frac{\psi_{sn}}{L_s - ML_m} \quad (11)$$

the expression of natural stator flux is given as

$$\psi_s = \psi_{n0} e^{-\frac{t}{\tau}} \quad (12)$$

where ψ_{n0} is the initial value of natural stator flux and $\tau_s = \frac{L_s - ML_m}{r_s}$ is the stator flux time constant with improved demagnetization control. It is clear that the damping of natural stator flux is accelerated with positive M. To have a small time constant, the coefficient M is tuned as large as possible. Therefore a demagnetizing method is obtained without depending on system parameters. The decreasing time constant can have the same effect compared to the traditional demagnetizing method, except that its dependence on system parameters is decreased. The effectiveness of this method is shown in the simulation results. Also in the simulation results section, the proposed method in this paper is compared with one of the best demagnetizing control methods in [22]. According to [22] forcing demagnetization control (FDC) modelling (shown in Fig. 3) is enhanced for LVRT capability of DFIG-based wind farm.

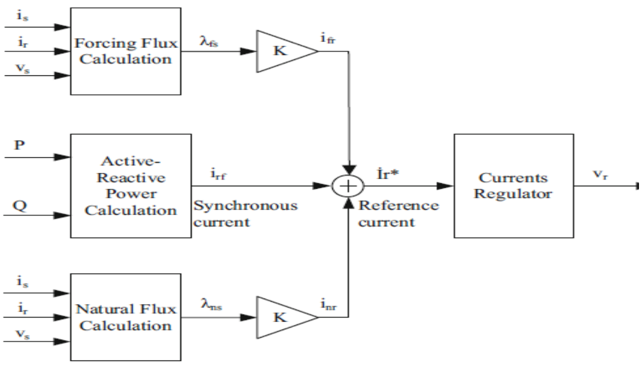


Fig.3. The block diagram of the proposed demagnetization control in [22].

3.2. Fault Identification

On the way to hastily switch the normal operation controller to the LVRT operation controller, quickly identifying of voltage dip may be very critical [26]. In [27] the fault occurrence time is identified through a suggested fuzzy logic controller. Another way for the identification of voltage dip in symmetric dips may be realized with the aid of the usage of transforming three-phase voltage variables in a natural reference frame to a synchronous d–q reference frame. The transformed d–q variables are proportional to the magnitude of the stator voltage; therefore, as compared to the values of d–q variables before voltage dip with ones after voltage dip, voltage dip may be recognized.

4. Simulation Results

The basic configuration of a DFIG-based wind turbine under the symmetric fault which is used in this paper is presented in Fig. 4. The understudy system in Fig. 4, is a wind farm which includes six 1.5 MW wind turbines. The three-phase fault depth is about 85 percentage and happens at second 1 with a duration of 0.3 S. The X/R ratio of the 1 Km cable is 3.4. The wind farm works in the maximum power point tracking (MPPT) mode with the 12 m/s wind speed. The grid frequency is 60 Hz and the DC link capacitor is 10000 micro Farad. Other DFIG parameters are presented in Table (1).

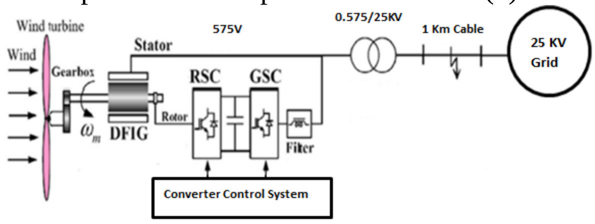


Fig.4. The general schematic of the DFIG-based wind turbine under the symmetric fault

In this part, the proposed demagnetizing method is simulated on the 9 MW wind farm and the results are compared to the proposed method in [22]. To see the effectiveness of the proposed method for improving the LVRT capability, the variations of the five parameters of stator terminal voltage (V_t), stator current (I_s), rotor current (I_r), rotor speed (ω_r) and dc-link voltage (V_{dc}), are analysed during the three-phase fault. The results are presented in Figs. (5-9).

Table 1. The DFIG parameters

Parameter	Values
Rated power	1.5 MW
Rated stator voltage	575 V
Rated frequency	60 Hz
Stator resistance	0.023 pu
Rotor resistance	0.016 pu
Stator leakage inductance	0.18 pu
Rotor leakage inductance	0.16 pu
Magnetizing inductance	2.9 pu
Pole pairs	3
Turns ratio	3.4

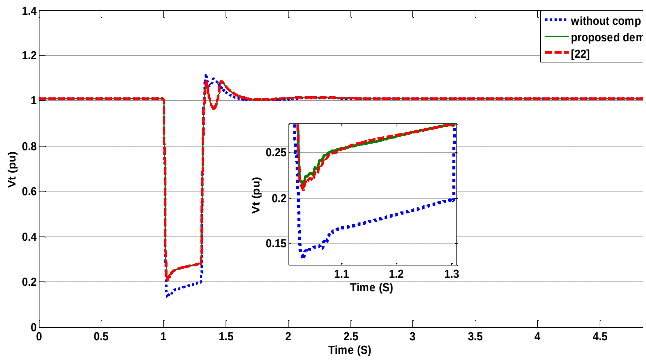


Fig.5. Stator voltage comparison

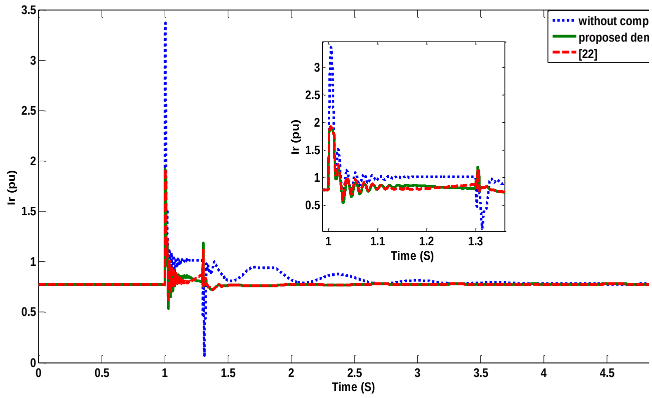


Fig.6. Rotor current comparison

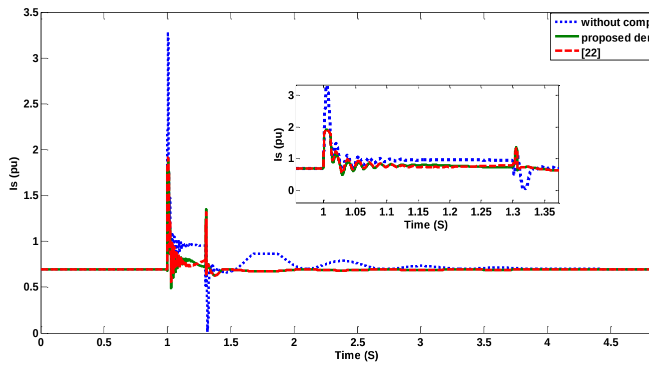


Fig.7. Stator current comparison

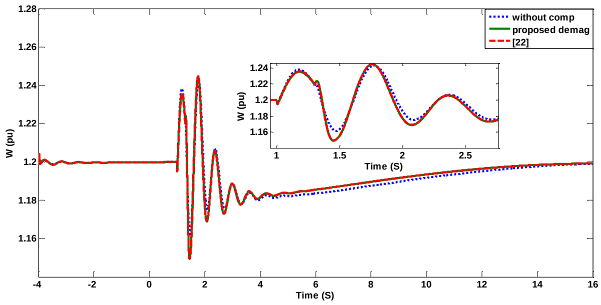


Fig.8. Rotor speed comparison

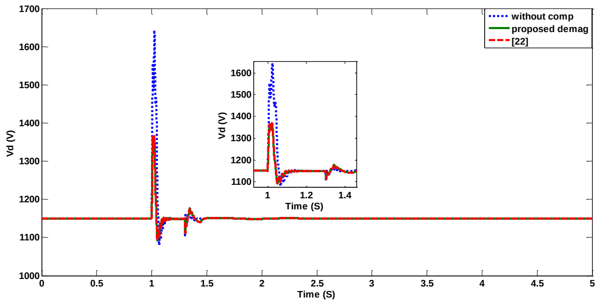


Fig.9. DC link voltage comparison

As seen in Fig. 5, according to LVRT criteria in Fig. 1, the LVRT capability is enhanced and compared to the proposed method in [22], the results are very similar. In Figs 6 and 7, the rotor and stator current is limited below 2.5 per unit, which is acceptable compared to the case which is without any demagnetizing control. In Fig. 8, although the rotor speed oscillations are more than the case without any demagnetizing control its oscillation range is acceptable. In Fig. 9, the peak of the DC-link voltage has decreased by about 20 percentage, which is below the danger zone. Despite acceptable results for improving LVRT capability, it is important to see if the proposed method works correctly in machine parameter changes or not. To check the dependence of the proposed demagnetizing control method on machine parameters, the effect of an increase in rotor resistance to 1.5 times greater (when the machine becomes warmer) and the decrease of magnetizing characteristics of machine to 0.5 times lower (due to machine ageing) is analysed in next part.

4.1. Rotor resistance increment

In this part, the effect of rotor resistance increase (to 1.5 times greater) on the performance of the proposed demagnetizing control method is analyzed and compared to the one proposed in [22]. The simulink results are presented in Fig. 10.

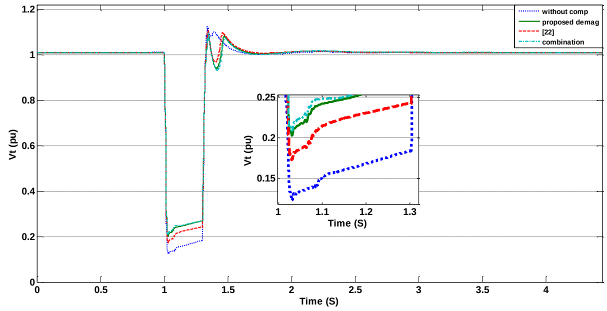


Fig. 10. Rotor resistance increase effect on stator voltage

As shown in Fig.10, the rotor resistance increase causes more voltage dip during fault duration. The voltage dip in the proposed method in [22] is more than 0.8 pu, which means that according to LVRT enhancement criteria in Fig. 1, a wind turbine must be disconnected from the grid. In contrast with the proposed method in [22], the voltage dip in the proposed method and the combination of the proposed method and the method proposed in [22], is less than 0.8 pu, which means that LVRT requirement is satisfied and improved. Therefore, it is proved that the dependency of this paper proposed method on rotor resistance increase is less than previous methods and still the LVRT capability is enhanced.

4.2. Magnetizing inductance decrement

In this part, the effect of magnetizing inductance decrease (to 0.5 times lower) due to machine ageing on performance of proposed demagnetizing control method is analyzed and compared to the one proposed in [22]. The simulink results are presented in Fig. 11.

As shown in Fig.11, the magnetizing inductance decrease causes more voltage dip during fault duration. The voltage dip in the proposed method in [22] is more than 0.8 pu, which means that according to LVRT enhancement criteria in Fig. 1, a wind turbine must be disconnected from the grid.

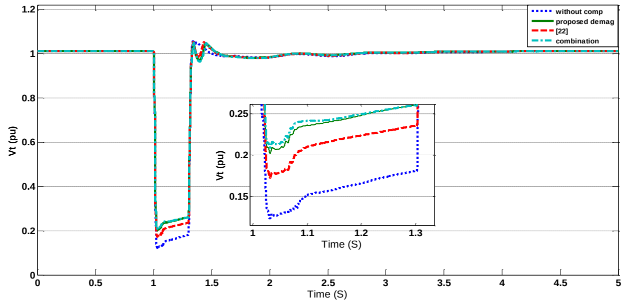


Fig. 11. Magnetizing inductance decrease effect on stator voltage

In contrast with the proposed method in [22], the voltage dip in the proposed method and the combination of the proposed method and the method proposed in [22], is less than 0.8 pu, which means that LVRT requirement is satisfied and improved. Therefore, it is proved that the dependency of this paper proposed method on magnetizing inductance decrease is less than previous methods and still the LVRT capability is enhanced.

4.3. Rotor resistance increment and magnetizing inductance decrement

According to Fig. 12, the effect of rotor resistance increase and simultaneously magnetizing inductance decrease on the performance of the proposed demagnetizing control method is analyzed and compared to the one proposed in [22].

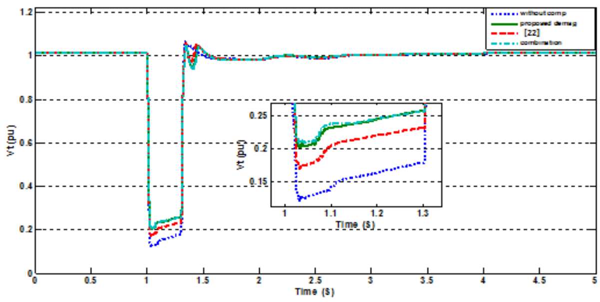


Fig. 12. Effect of increased rotor resistance and decreased magnetizing inductance on stator voltage

As shown in Fig.12, the two aforementioned machine parameter changes cause more voltage dip during fault duration compared to the last two cases. The voltage dip in the proposed method in [22] is more than 0.8 pu, which means that according to LVRT enhancement criteria in Fig. 1, a wind turbine must be disconnected from the grid.

In contrast with the proposed method in [22], the voltage dip in the proposed method and the combination of the proposed method and the method proposed in [22], is still less than 0.8 pu, which means that LVRT requirement is satisfied and improved. Therefore, it is proved that the dependency of this paper proposed method on simultaneously rotor resistance increase and magnetizing inductance decrease is less than previous methods and still the LVRT capability is enhanced.

5. Conclusion

In this paper, an enhanced LVRT control method on the rotor side Converter is proposed, which is based on demagnetizing natural stator current. The idea of this kind of demagnetizing control method is based on decreasing natural stator current time constant compared to decreasing natural stator flux magnitude. In this study, five important parameters which are stator voltage, rotor current, stator current, rotor speed, and DC link voltage were improved during a severe balanced three-phase fault and the results were analyzed. By using the proposed control method, the most important parameter which is the stator voltage was improved according to the LVRT criteria. The rotor and stator current and the dc-link voltage were improved too. Although the rotor speed was not improved through the proposed method, the oscillations are Limited. The previous methods were unable to enhance the LVRT capability in situations like rotor resistance increase or magnetizing inductance decrease. In contrast, the proposed method not only satisfies the LVRT requirement shown in Fig.1 but also has a good performance when the machine parameters change. Therefore the capability of the proposed method in this paper can enlarge the control range for deeper faults compared to the previous demagnetizing control methods.

References

[1] A. H. Kasem, E. F. El-Saadany, H. H. El-Tamaly, and M. A. A. Wahab, “A new fault ride-through strategy for doubly-fed wind-power induction generator,” in Proc. IEEE Elect. Power Conf., Canada, 2007, pp. 1–7.

[2] J. Lopez, E. Guba, P. Sanchis, E. Olea, J. Ruiz, and L. Marroyo, “Ride through of wind turbines with doubly-fed induction generator under symmetrical voltage dips,” IEEE Trans. Ind. Electron., vol. 56, no. 10, pp. 4246–4254, Oct. 2009.

[3] P. S. Flannery and G. Venkataramanan, “Fault-tolerant doubly-fed induction generator wind turbine using a parallel grid side rectifier and series grid side converter,” IEEE Trans. Power Electron., vol. 23, no. 3, pp. 1126–1135, May 2008.

[4] Grid Code, High and Extra High Voltage, E.ON Netz GmbH, Bayreuth, Germany, Apr. 2006. [Online]. Available: http://www.eon-netz.com

[5] W. Weisheng, C. Yongning, W. Zhen, et al., “On the Road to Wind Power: China’s Experience at Managing Disturbances with High Penetrations of Wind Generation,” IEEE Power Energy Mag., vol. 14, no. 6, pp. 24–34, Nov. 2016.

[6] G. Pannell, D.J. Atkinson and B. Zahawi, “Minim um-threshold crowbar for a fault-ride-through grid-code-compliant DFI G wind turbine”, IEEE Trans. on Energy Conversion, vol. 25, no. 2, pp. 750-759. Sept. 2010.

[7] Q. Wei, G. K. Venayagamorthy and R. G. Harley, “Real-time implementation of a STATCOM on a wind farm equipped with doubly fed induction generators,” IEEE Trans. Ind. Appl., vol. 45, no. 1, pp. 98– 107, Jan. 2009.

[8] P. S. Flannery and G. Venkataramanan, “A fault-tolerant doubly-fed induction generator wind turbine using a parallel grid side rectifier and series grid side converter,” IEEE Trans. Power Electron., vol. 23, no. 3, pp. 1126–1135, May 2008.

[9] P. S. Flannery and G. Venkataramanan. “Unbalanced voltage sag ride-through of a doubly-fed induction generator wind turbine with series grid-side converter,” IEEE Trans. Ind. Appl., vol. 45, no. 5, pp. 1879-1887, Sept. 2009.

[10] O. Abdel-Baqi and A. Nasiri, “ A Dynamic LVRT Solution for Doubly Fed Induction Generators”, IEEE Trans. Power Electro., vol. 25, no. 1, pp.193-196, Jan. 2010.

[11] C. Wessels, F. Gebhardt and F. W. Fuchs, “Fault ride-through of a DFIG wind turbine using a dynamic voltage restorer during symmetrical and asymmetrical grid faults”, IEEE Trans. on Power Electro., vol. 26, no. 3, pp. 807-815, March 2011.

[12] W. Guo, L. Xiao and S. Dai, “Enhancing low-voltage ride-through capability and smoothing output power of DFIG with a superconducting fault-current limiter–magnetic energy storage system”, IEEE Trans. on Energy Convers., vol. 27, no.2, pp. 277-295, June 2012.

[13] X. Yan, G. Venkataramanan, P. S. Flannery, Y. Wang, Q. Dong and B. Zhang, “Voltage-sag tolerance of DFIG wind turbine with a series grid side passive-impedance network”, IEEE Trans. on Energy Convers., vol. 25, no. 4, pp. 1048-1056, Dec. 2010.

[14] J. Yang, J. E. Fletcher and J. O’Reilly, “A series-dynamic-resistor-based converter protection scheme for doubly-fed induction generator during various fault conditions,” IEEE Trans. Energy Con vers., vol. 25, no. 2, pp. 422–432, Jun. 2010.

[15] L. Zhou, J. Liu, Y. Zhu and S. Zhou, “A Series Reactor Based Converter Protection Scheme of Doubly

Fed Induction Generator for Low Voltage Ride Through", IEEE Energy Conversion Congress and Exposition (ECCE), Raleigh, USA, pp. 2404-2409, Sept 2012.

[16] J. Liang, W. Qiao, and R. G. Harley, "Feed-Forward Transient Current Control for Low-Voltage Ride-Through Enhancement of DFIG Wind Turbines," IEEE Trans. Energy Convers., vol. 25, no. 3, pp. 836–843, 2010.

[17] J. Liang, D. F. Howard, J. A. Restrepo, and R. G. Harley, "Feedforward Transient Compensation Control for DFIG Wind Turbines During Both Balanced and Unbalanced Grid Disturbances," IEEE Trans. Ind. Appl., vol. 49, no. 3, pp. 1452–1463, 2013.

[18] D. Xiang, L. Ran, P. J. Tavner, and S. Yang, "Control of a doubly-fed induction generator in a wind turbine during grid fault ride-through," IEEE Trans. Energy Convers., vol. 21, no. 3, pp. 652–662, 2006.

[19] J. Lopez, P. Sanchis, E. Gubia, A. Ursua, L. Marroyo, and X. Roboam, "Control of Doubly Fed Induction Generator under symmetrical voltage dips," in 2008 IEEE International Symposium on Industrial Electronics, 2008, pp. 2456–2462.

[20] J. Lopez, E. Gubia, E. Olea, J. Ruiz, and L. Marroyo, "Ride through of wind turbines with doubly-fed induction generator under symmetrical voltage dips," IEEE Trans. Ind. Electron., vol. 56, no. 10, pp. 4246–4254, 2009.

[21] S. Hu, X. Lin, Y. Kang, and X. Zou, "An Improved Low-Voltage Ride-Through Control Strategy of Doubly Fed Induction Generator During Grid Faults," IEEE Trans. Power Electron., vol. 26, no. 12, pp. 3653–3665, 2011.

[22] M. K. Döşoğlu, U. Güvenç, Y. Sönmez, and C. Yılmaz, "Enhancement of demagnetization control for low-voltage ride-through capability in DFIG-based wind farm," Electr. Eng., vol. 100, no. 2, pp. 491–498, 2018.

[23] L. Yang, Z. Xu, J. Ostergaard, Z. Y. Dong, and K. P. Wong, "Advanced control strategy of DFIG wind turbines for power system fault ride through," IEEE Trans. Power Syst., vol. 27, no. 2, pp. 713–722, 2012.

[24] H. P. Darji, C. R. Mehta, and T. K. Tailor, "Comparative Analysis of Fault Current Limiters With and Without Shunt Path for LVRT Improvement of Grid Connected Wind Turbine Generator," in 2023 IEEE Renewable Energy and Sustainable E-Mobility Conference (RESEM), 2023, pp. 1–6. doi: 10.1109/RESEM57584.2023.10236258.

[25] B. Zhai, G. Chen, S. Alimasibieke, S. Liang, Z. Cao, and B. Gao, "Low Voltage Ride Through Method for Wind Turbines Based on Cooperative Control Strategy of Multiple Reactive Power Sources," in 2022 IEEE Sustainable Power and Energy Conference (iSPEC), 2022, pp. 1–6. doi: 10.1109/iSPEC54162.2022.10033017.

[26] D. Yazdani, M. Mojiri, A. Bakhshai, and G. Joos, "A fast and accurate synchronization technique for extraction of symmetrical components" IEEE Trans. Power Electron., vol. 24, no. 3, pp. 674–684, Mar. 2009.

[27] H. Rahimi Esfahani, A. Ketabi, H.R. mohammadi, and M. Rahimi, " Using VBR Model in Fixed Speed Wind Turbines and Suggesting a New Method for Improving LVRT Capability" Computational Intelligence in Electrical Engineering., vol. 10, no. 1, pp 51–62, Mar. 2019.

Journal of Optimization of Soft Computing (JOSC)

Journal of Optimization of Soft Computing (JOSC)

Vol. 1, Issue 1, pp: (39-53), September-2023

Journal homepage: https://sanad.iau.ir/journal/josc

A survey of meta-heuristic methods for optimization problems

Mehdi Fazli[1*] and Somayyeh Faraji Amoogin[2]

1,2. Department of Mathematics, Ardabil Branch, Islamic Azad University, Ardabil, Iran

Article Info

Article History:
Received 2023/10/6
Revised 2023/10/18
Accepted 2023/10/29

Keywords:
Operations research, Meta-heuristic, Hybridization, Memetic algorithm

**Corresponding Author's Email Address: mehdi.fazli.s@gmail.com*

Abstract

In this article, we will examine the problems related to routing and positioning with real variables and examine the related questions. These engineering, inventory and optimization decisions are made in a multi-layered supply chain system, including suppliers, warehouses and different buyers. We are looking for new ways to manage location and routing efficiently and effectively. In order to increase efficiency and achieve optimal results, exploratory and meta-heuristic methods have been used. In meta-heuristic techniques, a combination technique is usually used to increase performance. Therefore, this review article examines meta-heuristic methods and analysis of location problems using different quantities. It also examines the advantages and disadvantages of each method to optimally solve these problems in order to introduce practical and efficient methods.

1. Introduction

One of the practical and important issues in the field of optimization is the relationship between the set of facilities, possible warehouses and the set of applicants. The purpose of LRP is to provide better services to applicants and consider various facilities (planning routes and storage locations and cost limits). In these issues, the goal is to minimize the total cost of setting up warehouses, facilities, and turnover to serve each and every applicant. Facing such a problem, location and routing should be considered together and simultaneously, because neglecting any of the components in the location of facilities obviously increases the total cost of the distribution system and sometimes leads to a solution. unfavorable results (Salhi & Rand, 1989). From the perspective of complete supply chain management in the optimization discussion, location and routing seem to act as components of serious concern in real applications. As an urban application, the positioning of regional blood banks to serve patients Or and Pierskalla (1979), or establishing journal mail delivery systems

Jacobsen and Madsen (1980), or publishing goods parcels philanthropic–care network distribution dynasty Perl and Daskin (1985); Wasner and Zäpfel (2004)), to recitation a few, handle location and routing problems together Obviously, this problem-solving procedure will be challenging for logistics managers and decision-makers.

- In general, despite the effective application of the mechanisms and techniques of the meta-heuristic approach inspired by the real world for the problem of location allocation, according to scientific information, a general and systematic study in terms of antecedents and backgrounds has not been done in this regard. Therefore, in this study, we aim to review the existing real-world inspired algorithms for the specific problem of location allocation and routing, evaluate the differences of the mentioned techniques, and draw serious competitions and important issues about the mentioned problem that can be addressed in a wider field. In general, the contents and main concepts of this article can help in the following areas:

- Providing insights and implementations of location issue optimization techniques.
- To highlight the importance of using various methods of advanced meta-heuristic optimization, and the many advantages and disadvantages that they provide to deal with the challenges encountered in the location allocation problem.
- Determining relevant new and open issues and clues and ways to solve existing issues.
- In this paper, Section 2 provides an analysis and review of related work. Section 3 presents research terminology and preparation and implementation mechanisms. Section 4 discusses the review of selected meta-heuristic methods for different LRP problems. Section 5 presents the same new issues. Finally, Section 6 concludes the paper.

Table 1 shows the jointly used abbreviation in the article.

Table 1. Abbreviation table

Abbreviation	Definition	Abbreviation	Definition
LRP	Location Routing Problem	GA	Genetic algorithm
HM	Hybrid meta-heuristic	MA	Memetic algorithm
ACO	Ant colony optimization	PSO	Particle swarm optimization
HA	Hybrid Algorithm	HEA	Hybrid evolutionary algorithm
COA	Cuckoo optimization algorithm	TS	Tabu search
FA	Firefly algorithm	SA	Simulated annealing
SDO	Saturation degree-based ordering	HGS	Hybrid Genetic simulated annealing
MFFA	Memetic firefly algorithm	ACS	Ant colony system
APSO	Particle swarm optimization		

2.Related work

Some diverse studies have been conducted in the field of LRP. In this section, we review the literature for location routing conditions or how to use meta-heuristic algorithms in LRP and their strengths and weaknesses. This review was also conducted on articles from 2002 to 2022. A study of location routing problem is presented by Drexl and Schneider (2015) due to cost minimization.
One of the significant studies of the location-routing problems has been provided out by Prodhon and Prins (2014) . The LRP process integrates two major types of objectives. According to the existing and potential warehouses with variable costs, this transport fleet and the set of applicants have special demands. Classical LRP problems such as opening a subset of warehouses, assigning applicants to them, and specifying vehicle routes are used to minimize total costs such as specific storage costs and fixed transportation costs. They are also used throughout the project to minimize total costs. Since the last comprehensive scrutiny on LRP published by Nagy and Salhi (2007), The number of published articles devoted to this particular topic is rapidly increasing, and there is a need to analyze and review comprehensive new research works. This survey article analyzes the recent literature (68 articles) on standard LRPs and new additions such as access to multiple distributions and costs.
Vallada, Ruiz, and Minella (2008) Presented a review and extensive evaluation of heuristics and met heuristics for the m-machine stream shop scheduling problem aiming at minimizing general tardiness. He implemented 40various heuristics and met heuristics and examined their performance using the same criterion of instances. Their paper has presented an extensive and comprehensive review of heuristic and met heuristic procedures for the permutation flow department store scheduling problem aiming at minimizing the total tardiness. They have encrypted and tested 40 different algorithms to test their method performance under the same conditions. In addition, a benchmark is provided to evaluate all steps under a common dataset of 540 problems and a maximum of 350 related jobs and 50 new device.

Despite the important conditions of heuristic techniques fulfilled in the topic of localization, a global and systematic review of the discussion about their classification has not been presented so far. In addition, future challenges and the critical role of met heuristic techniques in a location problem are not well presented. In general, the reviewed articles have several reasonable categories as follows:

- ❖ The articles selection process has not been identified well.
- ❖ Some authentic and written researches have not been scrutinized.
- ❖ Future operations and open issues are not well stated.
- ❖ The categorization of the approaches studied has not been properly described.
- ❖ The strengths and weaknesses of the reviewed articles are well highlighted.
- ❖ In many studies, few papers have been reviewed.

3. Systematic literature review

The purpose of a review research is to provide a general and comprehensive summary of the literature related to the research domains related to the problem(Aznoli & Navimipour, 2017). Inspired by the field of medicine (Kitchenham, 2004), SLR, as a research procedure, provides a repeated procedure in which supplying sufficient details for being regurgitated by other researchers is carried out in(Charband & Navimipour, 2016; Cook et al., 1997; Kupiainen, Mäntylä, & Itkonen, 2015; Navimipour & Charband, 2016). A thorough search of the literature for relevant spread studies is the first step in conducting a principled review(Navimipour & Charband, 2016; Soltani & Navimipour, 2016).This has been supported by foregone studies suggesting that this procedure for literature review can lead to limited systematic errors, reduced chance factors, and enhanced validity of data analyses (Aznoli & Navimipour, 2016, 2017; Navimipour & Charband, 2016).

4. Review meta-heuristic algorithms in location problem

In this part of the article, we examine selected meta-heuristic methods and techniques from the wide range of studies in the optimization problem and especially location. Considering the role of the selection method in Section 3, we present various techniques in the groups of Particle Swarm Optimization (PSO), Ant Colony Optimization (ACO), Genetic Algorithm (GA) and Firefly Optimization (FA), Tabu Search (TS), We introduce and analyze the simulated annealing technique and the combined algorithm (HA) method. Qualitative measures for the location problem include runtime, comparison, real programs, competition, future development, algorithm improvement, parameter constraints, solution space reduction, and success.

4.1 Ant colony optimization method

Ant Colony Optimization is a collaborative intelligence-based search algorithm inspired by the behavior of real ants. First, it was practical and applied to solve the traveling salesman problem (TSP), and then it was effectively used to solve a large number of diverse problems such as routing in telecommunication networks, quadratic assignment problem (QAP), and graph coloring. This process is largely based on the idea of indirect communication among worker ant colonies. They go their way by borrowing Fermon. This pheromone pathway helps them find the shortest path between food sources and their nest. In the following, this process led to the formation of the mentioned method.

4.1.1 A summary of the selected process

A meta-heuristic approach, ACO, has suggested by S. Liu et al. (2017). In the proposed method, selection of pheromone models is an important priority. It is natural to choose pheromone models having fewer constraints; this algorithm can be used especially for the passenger seller problem in which the first order pheromone is widely recognizable. In the tests used, the algorithms using the second model are capable of general searching and have better population diversity. In fact, there is a qualitative difference between these two techniques. As a conclusion, they believe that different pheromone models can be used for problems on small and large scales, while the need to develop this method is extremely sensible.

An effective method has been developed by L. Yang et al (2013). to solve the facility location optimization problem with assumed single resource capacity (SSCFLP). This approach has been studied in three comparative studies of less than two different data size requirements, which can support the use of multiple computers to construct MAS. It also enhances the efficiency of smart computing. In numerical problems, data analysis based on practical examples showed that when multi-agent parallel calculations are performed on several systems, we reach intelligent calculations. In addition, the speed of the data network transmission process increases the computational efficiency of the multi-agent distribution method. The calculations and reproducibility of this method are high despite the favorable efficiency.

4.1.2. Some points of the ACO method that have been reviewed

Paper	Advantage	Disadvantage
S. Liu et al. (2017)	• Improve the existing solution • Application in practical matters • Providing promising solutions	• Development of the method in the future • Failure to check the same control parameters
L. Yang et al.(2013)	• High computing efficiency • Apply most restrictions • Logical calculation time • Cost Estimation	• Many repetitions • not always able to find the optimal solutions

Table 2 shows major advantages and disadvantages of ACO method for each research.

Table2. Major advantages and disadvantages of ACO method

4.2 Firefly algorithm (FA) optimization method

In this method and the corresponding algorithm, all fireflies are considered the same, regardless of the gender of the opposite fireflies, and they are attracted only based on the energy level of people. The firefly technique is produced based on the ideal guidance of the flashing characteristics of the firefly. In total, the following three rules can be used in the implementation of these features of the proposed method.

• All fireflies are born unisex, meaning a firefly can attract other things regardless of its gender.

• Their appeal is due to their brightness, that is, for both flashlights, the light travels less and decreases with increasing distance. If no one is brighter than a firefly, it moves randomly.

• The energy intensity of a firefly is influenced by its function and purpose and is optimized.

4.2.1 A summary of the selected process

A nature-inspired optimization method called firefly technique (FA) was proposed by Rao and Kumar (2014) to solve the multi-objective problem, which has multiple objectives based on a multi-criteria objective function. which tries to reduce the total power loss to a minimum, and tries to minimize the total voltage deviations. The total cost is equal to the fuel cost of producing and loading the branch and the optimal flow tries to reduce this cost. The results of this study show that the installation of SVC can increase the stability of the entire system voltage to a sufficient extent. In addition, the comparison between their proposed technique and a specific optimization method called genetic method showed that FA is an easy-to-use, robust and powerful optimization technique. In addition, the integration of SVC into the separate IEEE 14 and IEEE 30 system can reduce the total active power loss and improve the system performance. However, this method involves many iterations which can lead to longer execution time and is also designed for a specific problem.

Prima and Arymurthy (2019) considered the school transportation situation as a school location-allocation optimization problem. And with the help of optimization methods and algorithms based on nature, they tried to reduce the total distance walked by students to a favorable level and according to this process, they formulated the problem. This particular problem is assumed and modeled as a p-median problem, in which a meta-heuristic approach can be used. In this study, the firefly technique with 2D firefly representation is used to solve the location-assignment problem of the first school in South Jakarta. Therefore, gradual randomness reduction may improve the quality of the solution. Another improvement in the convergence of this technique is to change the parameter slowly, as the optimality is achieved and it can be useful for further research. In addition, as a relatively simple proposal, the firefly technique can be used to solve optimization in research on multi-objective and multi-component operations. However, for large scales, the method needs further development.

Suleiman et al. (2012) used the firefly technique to allocate the location and useful size of distributed generation (DG) in a convenient electricity distribution system. As already mentioned, FA is a meta-innovative technique related to the flashing behavior of fireflies. This behavior of the creatures is applied as a signal system in attracting other neighboring worms as the main target of the problem. In their paper, a new technique based on firefly movement in crowd is implemented to solve the problem of allocation and useful size of DG. Also, the efficiency and effectiveness of FA are discussed with some practical examples. The results show that the combination of DG in a distribution system can reduce the energy loss of the entire line and, as a result, significantly improve the system voltage characteristics. Also, compared to the performance of GA, the performance of FA is acceptable, in which GA also performs well in solving the optimal allocation problem. In this method, a distribution system with several DGs without DGs or with one DG and two DGs in the system was tested. The FA function is suitable for solving the problem of optimal location and size in the distribution system. In addition, adding DG to the distribution system can improve system performance, reduce total line losses, and improve load characteristics. However, the design of this method is not suitable for large-scale problems.

4.2.2. Some points of the FA method have been reviewed

Table 3 shows major advantages and disadvantages of FA method for each research.

Table3. Major advantages and disadvantages of FA method

Paper	Advantage	Disadvantage
Rao and Kumar (2014)	• Increase total power • High optimization power • Comparability	• Many repetitions Operating limit like time •
Prima and Arymurthy (2019)	• better results on medium scale • Comparability with other methods • Search space intensification • Increasing the performance of algorithms	• Need to develop algorithm • Unsuitable for real problems
Sulaiman et al. (2012)	• Improved voltage profile • Reduce line losses • Measuring distributed production • Optimal allocation of DG	• Need for future development • Suitable for small scale issues

4.3 Genetic Algorithm (GA) optimization method

In operational research and mathematics and computer science, the genetic algorithm and method is a meta-heuristic inspired by natural sciences that is used to select the next generation of people and is one of the evolutionary approaches. Genetic techniques are mainly used to generate high-efficiency solutions for optimization and search examples based on natural relationships. The main advantage of the genetic technique is the ability of this process to facilitate different subsets to evolve in different directions by applying other constraints. In examining different problems, it is proven that GA techniques improve the search process and can provide almost accurate and optimal solutions for different problems.

4.3.1 A summary of the selected process

Crossland et al. (2014) provided a useful tool for planning and solving heuristics using a simulated system for the problem in question. They use a genetic technique to investigate and solve the problem of electrical storage and measurement of energy storage in LV electrical systems. This particular method is used to check the configuration, storage topology and solve the electrical problem to solve the voltage increase problems as a result of increasing the efficiency of the PV system. Due to the electric absorption by PV, distributed storage is a good alternative to create and maintain the LV system. In addition, it is shown that a single-phase electrical configuration inside the main enclosure can solve the flow problem using a two-phase or three-phase model in the neighborhood. Therefore, this research provides a useful optimization tool for electrical positioning system in distributed LV networks. Single-phase storage at the applicant's site offers solutions that should be evaluated in terms of lower energy and three-phase storage at the main entrance. Home storage can benefit applicants by saving energy and reducing consumer bills due to the energy tariff structure. This generates more revenue to offset system costs. In addition, due to some limitations and attention to a specific case, this method needs to be developed.

Hiassat et al. (2017) used a mathematical model for location routing for perishable goods. In this problem model, the number and location of the warehouse of products throughout the year, the inventory of each seller and the shipping route are evaluated by all the variables of the problem. Their proposed model considers location allocation management decisions in a known inventory routing problem. Simultaneous consideration of these criteria and integration of strategic, tactical and operational management decisions will bring more useful results for the product supply chain. Considering that the model developed here is a hard problem and cannot find practical solutions in

a timely manner without providing a specific technique, they use a genetic technique to solve the problem efficiently and effectively. The obtained results prove that this particular model is more affordable than similar models. Possible formats include using several
different commodities, using different storage time windows for products and commodities (depending on time of year or temperature, etc.), and calculating carbon emissions, which the authors and scientists are working on. The genetic technique here effectively solves large and small samples. The solution of this model has been done using meta-heuristic method. Success is the result of the fact that the exact method does not solve many cases in a reasonable amount of time. To optimize the performance of this method, various parameters of the technique have been studied and tested. Despite the use of stochastic models and parameter constraints, the proposed method calculates the best solutions among similar methods.
Rybičková et al. (2016) stated and implemented a location routing problem. Where it was difficult to solve for NP-hard problems, they used meta-heuristics. This problem is used in supply chain optimization projects and distribution systems. The goal of solving the problem is routing, locating multiple warehouses, along with shipping routes, so that we reach a low amount in terms of costs. In the proposed problem of this research, a genetic method is presented to solve a routing problem. Individual representatives are designed with genetic problem operators to interpolate and solve both location and routing regions. Different parameters are presented using genetic technique to prove their effect on the results and acceptable performance of the proposed algorithm. The problem presented by them consists of optimal location of multiple warehouses along with optimal route distribution. The proposed algorithm continuously obtains better answers from the results of two similar genetic algorithms. In the experiment, several interesting dependences of the parameters on the results were found. Despite the dependence on parameters and the large volume of calculations, the proposed method can provide an optimal solution and reduce costs.

4.3.2. Some points of the GA method that have been reviewed

Table 4 shows major advantages and disadvantages of GA method for each research.

Table4. Major advantages and disadvantages of GA method

Paper	Advantage	Disadvantage
Crossland et al. (2014)	• Acceptable alternative in the real world • Real-world cost reduction • The best solution for random samples	• The need for real planning • Having some limitations • Improve the algorithm in the future
Hiassat et al. (2017)	• Comparability • Application in practical matters • Provide the best solution to similar methods	• Use random models • Parameter Limit
Rybičková et al. (2016)	• Minimizing the total cost of the system • Provide the optimal solution • Less computing time • reduction in costs • High sensitivity of the algorithm	• Dependence on parameters • Lot of calculations

4.4 Simulated Annealing (SA) method

The simulated annealing technique is used for problems in which many variables are used and we want to obtain a general optimum. In 1981, this method was used by researchers to find optimal solutions for problems. This process is often used when a discrete search space is considered. For problems in which finding an approximate global optimum is more important than finding a local optimum in a particular framework, simulated annealing may replace other similar techniques.

4.4.1 A summary of the selected process

Küçükoğlu et al. (2019) considered the process of uniform charging of electric vehicles at special stations for electric vehicles. They also considered that; Different technologies for batteries used in certain public stations lead to different battery charging times. Therefore, in this case study, the system (ETSPTW) is extended by considering the electrical operation in the applicant locations with different charge values. In this study, private and public electric stations for the system (ETSPTW) are modeled by applying other constraints. In addition to the developed version of ETSPTW, this paper also applies a new simulation technique (SA) to effectively and efficiently solve the ETSPTW-MCR problem. This algorithm uses efficient search methods based on system constraints (TSPTW), which uses a specific advanced solution acceptance process and planning for alternative solutions. However, the main purpose of this comparison is to show that the approach can be applied while respecting the quality of the solution.

Ghorbani and Jokar (2016) stated and formulated multi-stage and multi-product routing problems. There is a lot of competition in this field to solve the problem. For this paper, the simulated annealing method is implemented and a comprehensive numerical problem of multi-product routing is presented, which is investigated by the above method. In addition, the proposed method is solved and compared with the annealing technique in large and small samples. The results show that the proposed method and its algorithm in competition with (IC-SA algorithm) are better than other well-known algorithms in terms of time and CPU solution quality. In addition, the answers and solutions obtained using the IC-SA algorithm technique and the SA algorithm are evaluated in several small and large cases. The results obtained from the numerical examples show that: the IC-SA algorithm method performs significantly better than similar algorithms in terms of the quality of the obtained solution and the convergence of the solutions. However, trying to find a better and more effective algorithm is considered as a topic of future research. For example, a combination of simulated annealing technique (SA) and particle optimization technique (PSO) can be useful in comparison with IC-SA method. However, due to the high volume of calculations, the proposed method should be developed.

4.4.2. Some points of the SA method that have been reviewed

Table 5 shows major advantages and disadvantages of SA method for each research.

Table5. Major advantages and disadvantages of SA method

Paper	Advantage	Disadvantage
Küçükoğlu et al. (2019)	• solving routing problem simultaneously • Compatibility with other algorithms	• Long computing time • Many repetitions
Ghorbani and Jokar (2016)	• Convergent • Comparison with similar methods • high efficiency	• Many repetitions • Computation time

4.5 tabu search (TS) method

The root of the word taboo is taken from the Tongan language. Tonga natives and local people used them to represent things that could not be used because they were sacred. Forbidden search is a special technique that can be used to solve large and combinatorial optimization problems. In line with this method, in the absence of an improved move, worse moves can be accepted (such as when the search gets stuck at a local optimum). In addition, prohibitions (henceforth called taboos) are imposed to prevent previously visited solutions from returning. The implementation of Tabu search is a memory-based process. If you have visited a potential solution or violated a rule for a short period of time, you declare it "taboo" so that it will not be used again in the technique in question.

4.5.1 A summary of the selected process

Li et al. (2018) studied several objectives of LRP and compared it with several multi-objective meta-heuristic and heuristic approaches. In their problem, they have to place some plants and products in different environments to meet the demands of some applicants with the main goals of the project. This type of planning is formulated in management decisions to replace real data in the system. To solve this problem in a decision-making system, the metasearch technique based on tabu (TS) and its limitations were used. Using an adaptive memory method, MOAMP tries to adapt the tabu search method to the optimal setup structure of a multi-objective problem in the transportation network. Therefore, it takes an initial population of specified data (the first step of the technique) and uses it to obtain an optimal approximation of the rest using a neighborhood search process. This meta-heuristic approach addresses a real practical problem and provides information from two fire departments and an animal waste disposal plant in Andalusia, which obtains a broad approximation of the efficient set of problems. Future work includes trying to apply this pattern to non-neighboring points in larger areas. However, some new aspects of the metacognitive process, such as interactive mode and fuzzy information, will be used to provide sufficient information to the actual decision maker in the process. This method is suitable for small problems. In addition, solutions to larger problems must be formulated. In addition, the calculation time is not comparable with other methods.

Lai et al. (2016) investigated a heterogeneous time domain blocking problem for multiple objectives, where the parallax of different travel options is considered based on scales such as time, cost, distance, and number of trips. They expressed this problem as a complex numerical linear programming model in formula form, and proposed a tabu search exploration to efficiently solve the computational challenges of a parallel problem. In this case, numerical tests show that their proposed meta-algorithm is effective. Also, in the discussion of vehicle routing management, the main use of logistics and distribution systems has been. Responses indicate that there are significant savings in shipping costs by using an alternate route, especially when items are delivered to applicants within a limited time frame. This is because multiple structures use solutions using specific multiple arcs. For example, a special arc with higher travel costs but shorter travel time can be added to the vehicle route, which may reduce the need to dispatch an additional vehicle. Also, the large amount of calculation of this method increases its execution time.

Silvestrin and Ritt (2017) offered a variety of vehicle routing problem processing vehicles with different parts to get the best deal. Meta-heuristic methods are used to solve MCVRP problems. Inspired by signs (ITS), they search the location and provide the best solution. Accordingly, they find a solution by local search and then reach a local taboo by finding a taboo. Furthermore, they continue to search for the optimal value until they reach the stopping criteria. The need for different parts often occurs in practical applications. They have shown that ITS works well in VRP. In the existing methods, average quality and timing are applied in most cases. However, the methods may have general principles for different types of VRP. However, additional time did not significantly improve outcomes, whereas ITS improved significantly over time. The results of ITS surveys also show that the amount of various restrictions is relatively limited. This approach works for small problems, while it does not work well for large problems, so it needs to be developed for larger problems.

Polak and Boryczka (2019) presented a combination of location and vehicle decisions, which are known as NP-hard problems and require long computations and sufficient time. Due to the complexity of the problem, the discovery strategies simultaneously solved the routing problem and the routing decision. This two-dimensional architecture speeds up the search space, so that optimal solutions can be generated without excessive computation. An extensive study in this field shows that the performance of the TS algorithm is significantly improved over an active LRP problem. This method provides satisfactory answers to various problems and searches. However, in order to expand it, minor modifications should be made so that it can be effective for larger issues as well.

4.5.2. Some points of the TS method that have been reviewed

Table 6 shows major advantages and disadvantages of TS method for each research.

Table6. Major advantages and disadvantages of TS method

4.6 Particle Swarm Optimization (PSO) method

Paper	Advantage	Disadvantage
Li et al. (2018)	• Optimality more universal • Good computing time • Improve the performance of previous methods	• Run on small scales • improving algorithm in the future • no iterations • no runtime
Lai et al. (2016)	• System cost savings • Using the alternative path • Competing with other methods • Search space intensification	• Lot of calculations • Time of calculation • Approximation Method
Silvestrin and Ritt (2017)	• Competing with similar methods • Comparability • better results on medium scale	• Improve the method in the future • Lot of calculations
Polak and Boryczka (2019)	• Solving the two problems at the same time • Less computing • Low computing time	• Parameter Limit • Using Non-Optional Variables • Failure to adapt to real problems

In various sciences, the particle swarm optimization technique is an operational computational method that is optimized by repeatedly trying to refine a solution in observation with response measurement criteria. Consider this problem with a number of initial solutions, the particles considered as responses, and the movement of these particles into the search space. Researchers solve the problem by using simple mathematical formulas to move the particle in speed and position. At each step, the particles are directed to their known location as well as to known positions in the search space, which are updated when the system demands. This process is expected to lead to the best answer.

4.6.1 A summary of the selected process

Yu Peng et al. (2019) tested and presented a localization model with random parameters and artificial intelligence techniques. In their model, the location of fuzzy random possibility based on value at risk (VaR-FRFLM) is produced, in this model, both cost components and requests are assumed to be fuzzy random values, and the capacity of each possibility is fixed, and the value of the objective function is assumed to be continuous. A hybridization of the improved particle swarm optimization (MPSO) process is proposed to solve the VaR-FRFLM problem. In this developed technique, an approximation algorithm is used to calculate the stochastic fuzzy VaR target. The numerical results obtained show that the combination of MPSO performs better than when PSO and the corresponding regularized algorithm are used independently, but they only solve VaR-FRFLM. The capacity in the new model is not fixed, but it is assumed that the decision will be made. This action leads to the implementation of the model based on the running work, the proposed MPSO developed method and algorithm can be extended to other 0-1 integration optimization with VaR criterion. In addition, a VaR-based fuzzy stochastic temporal model may be far from a multi-objective model considering the expected profit as another objective. In general, the use of fuzzy variables is a growing result. Although there are optimization problems in random models, the effective efficiency of this method is acceptable.

Yang Peng and Chen (2009) presented a routing problem in the distribution network system and logistics operation, which mathematical example is presented in this research. Since finding the optimal solution for this example is a hard (non-polynomial) problem, they separate the main problem into several subproblems, for example one of the location allocation problems and the other of the car routing problem. Checks are done with the existing model and are applied with a penalty. The two-dimensional research is concerned with: (1) developing a model and method for multi-product routing and (2) combining the PSO method with another optimization algorithm to solve a more complex location problem than the routing problem. There are two problems at the same time and a suitable solution to the problem. However, this method suffers from parameter limitations despite providing suitable solutions for the intended problem.

Rabie et al. (2014) applied a particle swarm optimization (PSO) method to optimize large-scale continuous p-median location management decisions in the system. A related PSO technique, previously developed for the continuous P center, is used to optimize the p-mean position. Querying the location p is directed towards finding locations on the plane to minimize the total Euclidean distance from the range up to the set of final demand points. A notable aspect of the current work is that in the main topics, the problem was discussed in most of the cases already discussed in the literature. The obtained results are compared with several similar location problems. The Median-PSO-ED technique and method is used to find suitable solutions for problems with relatively large sizes. The results show the superiority of this technique compared to the results obtained in previous methods. In general, this method has been adapted to a specific problem and has yielded acceptable results, but it is not generalizable to other subjects.

Yu et al. (2014) proposed a location allocation problem for the location of urban parks based on the use of particle optimization (PSO) model. In the SLA model for public park, there are several operators such as competition, community compression and access. This shows that a number of parks have been investigated in this study. In order to find the required parking spaces, SLAs are checked using a very complex capping method. The PSO algorithm can reduce the computations and adjust the set of parks in the expected time. For example, by allocating places for hospital services, agriculture, water saving systems, cinemas and supermarkets, etc., this method can be used for other services. Compared to other artificial intelligence methods, this model is simpler and more practical and requires fewer parameters to calculate. The parameter of building size, the distance of networks from each other and local laboratory parameters for different places have an effect on the estimation of population density in the optimization method. The results obtained from SLA show that this particular approach is simple and cost-effective, and the system development tool can help local managers decide to do better work in urban planning. But in terms of computational time, this method has some shortcomings.

4.6.2. Some points of the PSO method that have been reviewed

Table 7 shows major advantages and disadvantages of PSO method for each research.

Table7 Major advantages and disadvantages of PSO method

Paper	Advantage	Disadvantage
Yu Peng et al. (2019)	• achieve the promising performance • Using Fuzzy Variables • Better time calculations	• Not generalizing the algorithm to integers • Random optimization problems • Not considering the expected profit
Yang Peng and Chen (2009)	• Solving the two problems at the same time • Provide practical solutions • Improve the performance of the algorithm	• Performance Limit
Rabie et al. (2014)	• Solving big problems • comparing results • Finding the best solution	• Having some limitations
Yu et al. (2014)	• Need less parameters • Can be generalized to most urban services in the real world • Simpler, more sophisticated than other AI techniques	• Computation time

4.7 Hybridization of meta-heuristic techniques

Currently, the hybrid of different techniques and the development of methods is one of the best and most prosperous trends in optimization. The combination of metacognition, such as evolution and optimization algorithms of ants and variable neighborhood search with artificial intelligence techniques and operations research, plays an important role. The resulting hybrid algorithms are usually labeled as meta-heuristic hybridization. The rise in this new research field is due to the fact that the basis of optimization research has changed from an algorithmic perspective to a general perspective. In this short review, we will examine hybrid methods and the use of fuzzy logic.

4.7.1 A summary of the selected process

(Zhang et al., 2022) consider the layout optimization problem with the desired complexities of the problem that the flow of materials and goods fluctuates between different parts within a certain range. A developed and mathematical model is used to reach the optimal solution. In order to achieve the overall goal of the project, which is MHC planning, the conditions of the floating material flow matrix are used in the production process. Considering that the new mathematical model introduces many intermediate parameters and variables, they are forced to use a multi-step solution method, and with this method, the time to solve the model is greatly increased. Next, they use a problem-based meta-heuristic optimization technique that develops and combines the advantages of the harmony search technique and the forbidden search technique, which integrates the solution phase of calling the exact solver into a two-step method. The project greatly improves facility layout and problem-solving efficiency. And it makes it possible to solve practical and real examples with this initiative. In their proposed model, the problem is implemented through a computer software, and then the mathematical model and the resulting combined algorithm are implemented together in the MATLAB software space. Finally, their proposed developed algorithm is used to solve practical and real problems that could not be solved by the two-step method before. In general, this method is effective for large and real problems.

(Lv et al., 2022) addressed a new location routing problem (LRP) by developing heuristics and called this form of optimization dual-mode integrated routing and location allocation (DMI) in a commodity supply chain logistics system. Warehouse products in their system meet the needs of online service requesters and retailers together, and customers and retailers follow two different order components. The problem in question includes several components that are presented in a general system. In that, the goods are referred from the manufacturing plant to the distribution centers (DC) in the first phase and from the centers (DC) to the buyers and retailers in the second phase. In this system, the mathematical model is presented along with some valid inequalities to strengthen the formula for the exact technique. They propose a solution for applicants and retailers in the problem of location allocation based on fuzzy correlation for each of the components. Following this process, a developed technique for neighbor point search (FCA) is presented for the planning and routing model. In their proposed method, tests have been performed on some standard samples and the results have been compared with some similar exploratory methods and other precise methods. The results of the examples and the optimization done show the effectiveness and feasibility of the proposed method, especially in solving real scale problems.

(Mokhtarzadeh et al., 2021) investigated the issue of new p-mobile hub allocation with all its components. In a system, hub facilities can be transferred to neighboring hubs for a new period. Implementing mobile hubs can contribute to hub opening and closing costs, especially in an environment where customer demands are rapidly increasing. On the other hand, moving the facilities of a network reduces the lifespan of the hubs and increases the related costs. Life expectancy and depreciation of hub facilities and other costs should be considered and specific restrictions on the movement of hub facilities should be applied. Several processes are in place to minimize the noise pollution, costs, and expense of creating a hub for the public. To optimize this proposed model, several meta-heuristic techniques, namely a non-dominated sorting genetic method (NSGA), multi-objective particle swarm optimization (MPSO), a combination of k-medoids as a well-known clustering method, and a combination of k-medoids and KMOPSO is used The obtained optimization results show that KNSGA-II has a good performance and is effective compared to other methods.

(Fazli et al., 2021) developed a two-phase hybrid heuristic method for solving the vehicle location problem (FLRP). In their methods, the FLRP model examines the location and routing objective. Upon arrival, they face a set of identical vehicles (each with fixed cost and capacity), a set of offices with limited opening cost, limited capacity, and a set of customers with fixed demand. In this case, they used fuzzy functions and variables to make the answers more realistic. The results and solutions of practical calculations show that the proposed mathematical model can solve other problems such as periodic routing problem (PLRP) and multiple routing problems (MDVRP) and several extensions of CLRP. In other words, adding constraints such as window time and heterogeneous fleets becomes problematic. In general, their proposed method is favorable in terms of calculation time compared to similar methods. With the help of using fuzzy values, it is effectively used in real problems.

(Saffari et al., 2022) developed an integrated logistics method in which the main decisions about storage location, route allocation by vehicle is made simultaneously. The balancing skill of joint decision criteria and total system costs shows that concurrent versions have capacity constraints in sequential problems, but they do not temporarily conflict with each other. Concurrent versions are also more effective non-dominant solutions than sequential versions. This procedure considers both problems at the same time and tries to reduce the cost of the system, so it has a challenging situation in terms of efficiency and timing of checks. They used the RBF-NN model by designing the fuzzy formula of the control parameters of the Whale Algorithm method. GWO, FWOA, WOA, GSA, ChOA, LCA, ES, GA algorithms have been used to use RBF-NN. The FWOA method can well define the boundary between exploitation and exploration phases. Therefore, this model does not get stuck in the local optimizer and proves its ability to find global optimizations to solve really high-dimensional problems. The obtained results show that RBF-WOA, RBF-FWOA, RBF-GWO, and RBF-ChOA methods have the best classification solutions, respectively. Also, the convergence speed of the FWOA method in reaching the optimal solution is better than the other seven used methods.

Babaie-Kafaki et al. (2016) have introduced a new method by hybridizing and increasing the number of iterations of genetic algorithms and gradually increasing the probability of using the neighborhood search method on specific members. They implemented the proposed hybrid technique and compared the performance of their method with several well-developed meta-heuristic techniques. In this context, they use a special method based on neighborhood search for all investigated points, in this process they apply two new simulated hybridization techniques to the best individual in the population in each iteration. They applied algorithms on the location problem using fuzzy values, these fuzzy values were used to reduce uncertainty and make the problems more realistic. To investigate the effectiveness of the proposed technique, they implemented and solved specific models for the fuzzy terminal location problem. It is assumed that the fuzzy model has the number of terminals and passengers with fuzzy values corresponding to the hypothesis, as well as the fuzzy location with default upper and lower bounds for the number of locations required by the problem. They randomly generated algorithms at different terminal locations with large-scale fuzzy values, where the coefficient of the cost function is considered as a fuzzy value. Although the proposed method has many iterations, it is suitable for large and real problems.

Saif-Eddine et al. (2019) introduced two new hybrid techniques. They solve the optimization problem by combining existing methods and forming a new hybrid method (HGA). Their paper deals with the inventory location allocation and routing problem (ILRP) when making managerial decisions. Considering the constraints and to minimize the cost of the whole supply chain, a mathematical model has been used and this is a hard problem. A new improved genetic algorithm (IGA) is designed and modeled to solve the real-scale problem optimally. In the implemented problems, two samples (30 and 10 applicants) have been planned and solved. To demonstrate the impact of this new approach, the total vehicle capacity (number of vehicles in the parking lot and capacity of each vehicle) on the total cost of the supply chain is considered. In the examples, the obtained results show that the IGA technique performs better than the GA technique in achieving a lower cost, especially for a large number of applicants. Improving the efficiency of the solution and the obtained solutions is mainly achieved at the cost of optimal computing time. Despite the limitations of the parameters, this method is highly efficient and has a good computing time compared to similar methods.

4.7.2. Some points of the HA method that have been reviewed

Table 8 shows major advantages and disadvantages of HA method for each research

Table8 Major advantages and disadvantages of HA method

Paper	Advantage	Disadvantage
Zhang et al. (2022)	• Solving the two problems at the same time • Provide practical solutions • Improve the performance of the algorithm	• Having limits on big problems
Lv et al.(2022)	• Generate synchronous version • Average cost improvement • Low computing time • Average and proper distance	• Applying medium graphs
Mokhtarzadeh et al. (2021)	• Improve the quality and time of the CPU • Solving the two problems at the same time • Application for large samples	• Computation time
Fazli et al. (2021)	• Low computing • Using real values • Comparability	• Having some limitations
Saffari et al.(2022)	• Generate synchronous version • Average cost improvement • Low computing time • Average and proper distance	• Applying medium graphs

Babaie-Kafaki et al. (2016)	• Real-world application • Low computing time • It works for great issues	• Repeated algorithms
Saif-Eddine et al. (2019)	• Low computing time • Compare with similar algorithms • Little repeats	• Having some limitations • Using random data in the algorithm

5. Open issue

In this section, we will examine the different methods of the location allocation problem, which have not yet been widely and comprehensively addressed. This can be taken into consideration in Moore's future research and be a research guide. According to the meta-heuristic techniques used, in the location and routing problem, the priorities and criteria mentioned in this article are the number of repetitions, competition, and execution time, while some researches completely ignore these important issues. For example, in research, it is necessary to consider convergence algorithms, appropriate execution time, number of iterations and method improvement. Also, in some related techniques and algorithms, success, competition and comparison are considered, while these issues are not prioritized in some studies. Therefore, meta-heuristic techniques and their algorithms were analyzed in detail. There is no independent technique that addresses all criteria in a given problem. Another important point of future study can be the investigation of solving the location problem with the development of meta-heuristic algorithms. In addition, some meta-heuristic algorithms that are efficient and effective in specific subjects are considered. It will be interesting to examine their methods. When we increase the size of the location problem, it is clear that the number of iterations of the algorithm and the execution time increase. So that these problems do not occur, meta-heuristic methods and algorithms should be redesigned with the aim of eliminating the problems caused by such cases. In addition, the efficiency of meta-heuristic algorithms in solving real location problems is one of the important aspects, including in larger-scale problems. Therefore, the efficiency of the methods and their algorithms for convergent solutions, global optimality, number of successes, etc. are interesting and good criteria for future studies.

In most of the reviewed techniques, many new meta-heuristic algorithms such as particle swarm optimization, artificial fish algorithm, simulated annealing, etc. are used with the aim of increasing the efficiency for this special problem in prominent journals. In addition, some factors such as dependence of algorithms on parameters, starting with infeasible solutions, large scale graph, reduction of solution space and standard deviations were not considered. In addition, in this review article, the application of meta-heuristic algorithms in the location problem is fully discussed. By reviewing the studies, we observed that in hybrid meta-heuristic algorithms, better results are obtained than independent methods. In other words, most scientific criteria are improved by specific hybrid methods. Therefore, some new hybrid approaches are needed in future projects.

6. Conclusion

Location allocation and routing problems can have different effects on the economy of the society. Although meta-heuristic algorithms are practical methods for solving location problems, due to the computational time and more iterations, in recent years, the tendency to combine these methods has been suggested to increase their efficiency. Therefore, we propose a hybrid of these methods to solve these types of problems. These methods give the best results and at the same time work best on a set of problems. Again, the calculations show that: meta combination results are better than heuristic and metacognitive combinations. As its creation mechanism for generating new solutions, it allows storing and using relatively large equations of good and varied solutions during the search process. These methods require minimal computing time and are very suitable in line with similar experiments presented in the literature. In general, combined metacognitive methods are very effective tools for finding suitable solutions. Case studies illustrate the approaches needed for social learning research, as long-term process analysis requires sensitivity to social and economic contexts. Therefore, fuzzy data can be used for problem solving in obtaining a dataset of real problems. This issue can also be the subject of future research.

References

Amaya, J. E., Porras, C. C., & Leiva, A. J. F. (2015). Memetic and hybrid evolutionary algorithms. In *Springer Handbook of Computational Intelligence* (pp. 1047-1060): Springer.

Aznoli, F., & Navimipour, N. J. (2016). Deployment Strategies in the Wireless Sensor Networks: Systematic Literature Review, Classification, and Current Trends. *Wireless Personal Communications*, 1-28.

Aznoli, F., & Navimipour, N. J. (2017). Cloud services recommendation: Reviewing the recent advances and suggesting the future research directions. *Journal of Network and Computer Applications, 77*, 73-86.

Babaie-Kafaki, S., Ghanbari, R., & Mahdavi-Amiri, N. (2016). Hybridizations of genetic algorithms and neighborhood search metaheuristics for fuzzy bus terminal location problems. *Applied Soft Computing, 46*, 220-229.

Charband, Y., & Navimipour, N. J. (2016). Online knowledge sharing mechanisms: a systematic review of the state of the art literature and recommendations for future research. *Information Systems Frontiers, 6*(18), 1131-1151.

Crossland, A., Jones, D., & Wade, N. (2014). Planning the location and rating of distributed energy storage in LV networks using a genetic algorithm with simulated annealing. *International Journal of Electrical Power & Energy Systems, 59*, 103-110.

Drexl, M., & Schneider, M. (2015). A survey of variants and extensions of the location-routing problem. *European Journal of Operational Research, 241*(2), 283-308.

Ellis, R., & Petridis, M. (2009). *Research and Development in Intelligent Systems XXVI: Incorporating Applications and Innovations in Intelligent Systems XVII*: Springer Science & Business Media.

Fazli, M., F.M.Khiabani and B. Daneshian, *Hybrid whale and genetic algorithms with fuzzy values to solve the location problem* mmep, 763-768

Ghorbani, A., & Jokar, M. R. A. (2016). A hybrid imperialist competitive-simulated annealing algorithm for a multisource multi-product location-routing-inventory problem. *Computers & Industrial Engineering, 101*, 116-127.

Hiassat, A., Diabat, A., & Rahwan, I. (2017). A genetic algorithm approach for location-inventory-routing problem with perishable products. *Journal of Manufacturing Systems, 42*, 93-103.

Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele, UK, Keele University, 33*(2004), 1-26.

Küçükoğlu, İ., Dewil, R., & Cattrysse, D. (2019). Hybrid simulated annealing and tabu search method for the electric travelling salesman problem with time windows and mixed charging rates. *Expert Systems with Applications, 134*, 279-303.

Kupiainen, E., Mäntylä, M. V., & Itkonen, J. (2015). Using metrics in Agile and Lean Software Development–A systematic literature review of industrial studies. *Information and Software Technology, 62*, 143-163.

Lai, D. S., Demirag, O. C., & Leung, J. M. (2016). A tabu search heuristic for the heterogeneous vehicle routing problem on a multigraph. *Transportation Research Part E: Logistics and Transportation Review, 86*, 32-52.

Li, X., Yue, C., Aneja, Y. P., Chen, S., & Cui, Y. (2018). An Iterated Tabu Search Metaheuristic for the Regenerator Location Problem. *Applied Soft Computing, 70*, 182-194.

Liu, S., Leng, H., & Han, L. (2017). Pheromone Model Selection in Ant Colony Optimization for the Travelling Salesman Problem. *Chinese Journal of Electronics, 26*(2), 223-229.

Lv, C., Zhang, C., Ren, Y., & Meng, L. (2022). A fuzzy correlation based heuristic for Dual-mode integrated Location routing problem. *Computers & Operations Research, 146*, 105923.

Mokhtarzadeh, M., Tavakkoli-Moghaddam, R., Triki, C., & Rahimi, Y. (2021). A hybrid of clustering and meta-heuristic algorithms to solve a p-mobile hub location–allocation problem with the depreciation cost of hub facilities. *Engineering Applications of Artificial Intelligence, 98*, 104121.

Nagy, G., & Salhi, S. (2007). Location-routing: Issues, models and methods. *European Journal of Operational Research, 177*(2), 649-672.

Navimipour, N. J., & Charband, Y. (2016). Knowledge sharing mechanisms and techniques in project teams: literature review, classification, and current trends. *Computers in Human Behavior, 62*, 730-742.

Peng, Y., & Chen, Z.-X. (2009). *Two-phase particle swarm optimization for multi-depot location-routing problem.* Paper presented at the New Trends in Information and Service Science, 2009. NISS'09. International Conference on.

Peng, Y., Cheng, J.-f., & Jiang, R.-x. (2019). Inversion of UEP signatures induced by ships based on PSO method. *Defence Technology*.

Perl, J., & Daskin, M. S. (1985). A warehouse location-routing problem. *Transportation Research Part B: Methodological, 19*(5), 381-396.

Polak, I., & Boryczka, M. (2019). Tabu Search in revealing the internal state of RC4+ cipher. *Applied Soft Computing, 77*, 509-519.

Prima, P., & Arymurthy, A. M. (2019). *Optimization of school location-allocation using Firefly Algorithm.* Paper presented at the Journal of Physics: Conference Series.

Prodhon, C., & Prins, C. (2014). A survey of recent research on location-routing problems. *European Journal of Operational Research, 238*(1), 1-17.

Rabie, H. M., El-Khodary, I. A., & Tharwat, A. A. (2014). *Particle Swarm Optimization algorithm for the continuous p-median location problems.* Paper presented at the Computer Engineering Conference (ICENCO), 2014 10th International.

Rao, B. V., & Kumar, G. N. (2014). Sensitivity analysis based optimal location and tuning of static VAR compensator using firefly algorithm. *Indian Journal of Science and Technology, 7*(8), 1201-1210.

Rybičková, A., Burketová, A., & Mocková, D. (2016). *Solution to the location-routing problem using a genetic algorithm.* Paper presented at the Smart Cities Symposium Prague (SCSP), 2016.

Saffari, A., Zahiri, S. H., & Khishe, M. (2022). Fuzzy whale optimisation algorithm: a new hybrid approach for automatic sonar target recognition. *Journal of Experimental & Theoretical Artificial Intelligence*, 1-17.

Saif-Eddine, A. S., El-Beheiry, M. M., & El-Kharbotly, A. K. (2019). An improved genetic algorithm for optimizing total supply chain cost in inventory location routing problem. *Ain Shams Engineering Journal, 10*(1), 63-76.

Salhi, S., & Rand, G. K. (1989). The effect of ignoring routes when locating depots. *European journal of operational research*, *39*(2), 150-156.

Silvestrin, P. V., & Ritt, M. (2017). An iterated tabu search for the multi-compartment vehicle routing problem. *Computers & Operations Research, 81*, 192-202.

Soltani, Z., & Navimipour, N. J. (2016). Customer relationship management mechanisms: A systematic review of the state of the art literature and recommendations for future research. *Computers in Human Behavior, 61*, 667-688.

Sulaiman, M. H., Mustafa, M. W., Azmi, A., Aliman, O., & Rahim, S. A. (2012). *Optimal allocation and sizing of distributed generation in distribution system via firefly algorithm.* Paper presented at the Power Engineering and Optimization Conference (PEDCO) Melaka, Malaysia, 2012 Ieee International.

Vallada, E., Ruiz, R., & Minella, G. (2008). Minimising total tardiness in the m-machine flowshop problem: A review and evaluation of heuristics and metaheuristics. *Computers & Operations Research, 35*(4), 1350-1373.

Wasner, M., & Zäpfel, G. (2004). An integrated multi-depot hub-location vehicle routing model for network planning of parcel service. *International Journal of Production Economics, 90*(3), 403-419.

Yang, L., Sun, X., & Chi, T. (2013). *An ant colony optimization algorithm and multi-agent system combined method to solve Single Source Capacitated Facility Location Problem.* Paper presented at the Advanced Computational Intelligence (ICACI), 2013 Sixth International Conference on.

Yu, J., Chen, Y., Wu, J., Liu, R., Xu, H., Yao, D., & Fu, J. (2014). *Particle swarm optimization based spatial location allocation of urban parks—A case study in Baoshan District, Shanghai, China.* Paper presented at the Agro-geoinformatics (Agro-geoinformatics 2014), Third International Conference on.

Zhang, Z., Gong, J., Liu, J., & Chen, F. (2022). A fast two-stage hybrid meta-heuristic algorithm for robust corridor allocation problem. *Advanced Engineering Informatics, 53*, 101700.