فصل نامه تخصصی مهندسی مخابرات جنوب
دانشگاه آزاد اسلامی واحد بوشهر

فصل نامه تخصصی مهندسی مخابرات جنوب

صاحب امتیاز : دانشگاه آزاد اسلامی واحد بوشهر

نشانی : بوشهر، دانشگاه آزاد اسلامی واحد بوشهر، حوزه معاونت پژوهشی، دفتر مجله تخصصی مهندسی مخابرات جنوب

تلفن : ۰۹۱۰۷۸۳۷٤۲۰ **فاکس** :۰۷۷۳۳۵۵۵٤۱۳

سایت نشریه: https://jce.bushehr.iau.ir **شاپا الکترونیکی**: ۲۹۸۰-۹۲۳۱

نشانی الکترونیکی : jce.iaub@gmail.com - jce@iaubusher.ac.ir

نشریه در پایگاههای ملی و بین المللی زیر نمایه شده است:

DOAJ DIRECTORY OF OPEN ACCESS JOURNALS | معاونت پژوهشی دفتر سیاست گذاری و برنامه ریزی امور پژوهشی | magiran | پایگاه اطلاعات علمی جهاد دانشگاهی SID

Google Scholar | پایگاه استنادی علوم جهان اسلام ISC | CIVILICA We Respect the Science | ACADEMIA

این مجله بر اساس مجوز انتشارشماره ۸۷/٤۲۲۰۹۲ تاریخ ۸۹/۱۰/۲۲ مدیر کل دفتر گسترش تولید علم دانشگاه آزاد اسلامی انتشار می یابد.

فصلنامه تخصصی مهندسی مخابرات جنوب
دانشگاه آزاد اسلامی واحد بوشهر
شاپا: ۹۲۳۱-۲۹۸۰
سال چهاردهم، شماره پنجاه و شش، تابستان ۱۴۰۴

فهرست

JOURNAL OF SOUTHERN COMMUNICATION ENGINEERING
ISLAMIC AZAD UNIVERSITY BUSHEHR BRANCH

E-ISSN: 2980-9231
https://sanad.iau.ir/journal/jce

https://doi.org/10.30495/jce.2024.2003925.1242

Vol. 14/ No. 56/2025

Research Article

A New DNA Cryptosystem for Encrypting Persian Texts and Images with Key Exchange based on Hyper Elliptic Curve

Fatemeh Alidadi Shamsabadi[1] | Shaghayegh Bakhtiari Chehelcheshmeh[2]* | Narges Farhadi[3]

[1]Department of Computer, ShK.C., Islamic Azad University, Shahrekord, Iran.
fa.alidadi@iau.ir

[2]Department of Computer, ShK.C., Islamic Azad University, Shahrekord, Iran.
bakhtiari@iau.ac.ir

[3]Department of Computer, ShK.C., Islamic Azad University, Shahrekord, Iran.
n.farhadi562@iau.ac.ir

Correspondence
Shaghayegh Bakhtiari Chehelcheshmeh, Assistant Professor, Department of Computer, ShK.C., Islamic Azad University, Shahrekord, Iran.
bakhtiari@iau.ac.ir

Main Subjects:
DNA Cryptography

Paper History
Received: 16 December 2023
Revised: 28 January 2024
Accepted: 31 January 2024

Abstract

With the emergence of quantum computers, traditional cryptographic methods will be compromised and lose their effectiveness. Fortunately, DNA-based cryptographic methods exhibit high resistance against quantum attacks. However, these methods have not yet been employed for securing Persian texts. Additionally, DNA encryption methods require a secure communication channel for sharing the secret key. Therefore, this paper presents a new cryptosystem to ensure the confidentiality of Persian texts based on DNA cryptography. Furthermore, considering the widespread use of images in public media, the capability to encrypt images has also been incorporated into this scheme. In the proposed method, the Hyperelliptic Curve Cryptography (HECC) system is utilized to generate a shared secret key between communicating parties, eliminating the need for a secure channel. DNA hybridization technology is also employed for message encryption and decryption to minimize the computational complexity of the cryptographic algorithm. Performance evaluation results demonstrate that the proposed scheme reduces computational overhead and exhibits lower energy consumption compared to previous approaches.

Keywords: Security, HEC Key Agreement, DNA Cryptography, Persian Plaintexts, DNA Hybridization.

Highlights

- Encryption of Persian texts using a DNA cryptography system for the first time.
- Providing a large key space to increase resistance against external factors.
- Minimizing time complexity using DNA hybridization technology.
- No need for a secure channel for key sharing.
- Highly secure and efficient key agreement using encryption based on hyper-elliptic curve.

Citation: F. Alidadi Shamsabadi, Sh. Bakhtiari Chehelcheshmeh, and N. Farhadi, "A New DNA Cryptosystem for Encrypting Persian Texts and Images with Key Exchange based on Hyper Elliptic Curve," *Journal of Southern Communication Engineering*, vol. 14, no. 56, pp. 1–26, 2025, doi:10.30495/jce.2024.2003925.1242 [in Persian].

1. Introduction

Today, the transfer of digital data—especially text and image files—has increased significantly across computer networks. Ensuring data security and protecting it against unauthorized access remains a major challenge in these networks [1]. Encryption enables secure data transfer by storing information securely in channels, ensuring that only the intended recipient can decipher the transmitted data [2].

Various methods for encrypting information have been developed. One such method is symmetric cryptography, which performs encryption and decryption rapidly using the same secret key. However, this method requires a secure communication channel to share the secret key between both parties. To address this limitation, asymmetric cryptography employs a pair of keys: a public key for encryption and a private key for decryption. The public key can be shared openly without a secure channel, while the private key remains confidential [1,3]. This eliminates the need to exchange a secret key beforehand. However, public-key cryptography methods are less efficient (slower) [4].

For securely transferring large volumes of data, users typically first establish communication using asymmetric key cryptography. They then exchange a symmetric secret key and switch to symmetric encryption, thereby leveraging the advantages of both methods [5].

Unfortunately, traditional cryptographic algorithms lack sufficient computational power and storage capacity and are vulnerable to quantum attacks. As a result, DNA-based cryptographic methods are emerging as a promising solution for secure communication [2].

DNA cryptography bridges two fields: biology and computer science [6]. This method exploits the natural processes of DNA formation for data encryption [7]. Due to the high information density in DNA molecules, vast amounts of encoded data can be stored in a compact space. Additionally, DNA cryptography benefits from parallel computation, enabling extremely fast processing [8-11].

DNA cryptography integrates the massive parallelization and storage capacity of DNA molecules with traditional cryptographic methods (symmetric and asymmetric) [5]. However, a persistent challenge in DNA encryption systems is the need for a secure channel or an efficient key-agreement solution.

Elliptic curve cryptography is a key-agreement system known for its high speed and suitability for resource-constrained mobile nodes, offering enhanced security and energy efficiency. Compared to other cryptographic methods (e.g., RSA, DES, ECC), it provides superior operational efficiency and lower computational overhead. HECC (Hyperelliptic Curve Cryptography) is particularly versatile, allowing easy implementation in both hardware and software applications [12].

2. Innovation and contributions

This paper presents a novel scheme for encrypting and decrypting Persian texts and images using HECC (Hyperelliptic Curve Cryptography) combined with DNA cryptography, ensuring robust data security while minimizing computational overhead.

Key innovations of this work include:

- **Encryption of Images and Persian Texts Based on DNA Cryptography:** For the first time, an algorithm is presented for encrypting images and Persian texts using the DNA cryptography system. If the data is in text form, it is easily encrypted using this algorithm. However, if the data is in image form, the image is first converted into a DNA sequence. Then, the same text encryption method is used to encrypt it. Interestingly, the data obtained from the image is encrypted in the format of a text file. Therefore, the proposed scheme has an additional feature for hiding images as text files.
- **Use of DNA Hybridization Technology:** DNA hybridization is used to minimize time complexity in the proposed system. In the hybridization method, two single-stranded DNAs containing encrypted information are combined to produce a single DNA sequence.
- **Key Agreement Based on Hyperelliptic Curve:** In this scheme, HECC is used for secret key exchange. Therefore, there is no need for a secure channel to share the secret key. Additionally, the proposed key distribution has low computational overhead and high execution speed.
- **Providing a Large Key Space:** An efficient cryptography system should have a key that is large enough to ensure its security. Therefore, the proposed scheme offers a larger key space (640-bit) compared to other similar schemes.

3. Materials and Methods

The proposed scheme employs rigorous mathematical proofs, logical derivations, and practical implementation through coding in the Visual Studio development environment to validate its theoretical foundations and functional performance.

4. Results and Discussion

The proposed cryptosystem enhances security through two key mechanisms: (1) leveraging the inherent complexity of biological processes in DNA cryptography, and (2) employing computationally intensive

hyperelliptic curve operations for shared key generation. This dual approach enables secure encryption of both Persian texts and images without requiring a pre-established secure channel for key exchange.
The system's architecture combines the efficiency of symmetric cryptography with the security advantages of asymmetric methods, specifically through:

- DNA-based symmetric encryption of content
- Hyperelliptic curve asymmetric key establishment

Implementation results demonstrate significant improvements in:

- Computational efficiency (reduced overhead)
- Temporal performance (lower time complexity)
- Security robustness

These benefits derive from the synergistic integration of:

- DNA hybridization techniques
- HECC-based key agreement protocols

The mathematical soundness of the algorithm has been formally verified, with practical implementation confirming its theoretical advantages for secure data transmission scenarios.

5. Conclusion

DNA cryptography represents an innovative approach for message encryption that utilizes DNA molecules as information carriers. This method capitalizes on DNA's inherent advantages, including exceptional storage capacity, minimal energy requirements, and parallel processing capabilities, to facilitate secure encryption and decryption operations. The DNA-based approach enhances data transmission security across networks while demonstrating resistance against quantum computing attacks.
DNA cryptosystems exhibit compatibility with both symmetric and asymmetric encryption paradigms. A significant efficiency improvement is achieved by implementing shared key generation mechanisms that eliminate secure channel key transmission requirements. This paper introduces a novel DNA cryptosystem specifically designed for Persian text and image encryption, incorporating two key innovations:

1. HEC Cryptography for Key Generation: The system employs hyperelliptic curve cryptography to establish shared secret keys securely without channel dependency.
2. DNA Hybridization Implementation: The integration of DNA hybridization technology ensures rapid execution speeds while maintaining cryptographic robustness.

The proposed scheme offers enhanced security through an expanded cryptographic key space. Comprehensive security analyses and performance evaluations confirm the algorithm's superior precision, execution speed, and resistance to attacks compared to conventional encryption systems. These improvements stem from the synergistic combination of DNA molecular computing and advanced cryptographic techniques.

6. Acknowledgement

We sincerely thank the honorable reviewers for their valuable comments and suggestions that significantly improved the quality of this article. We also extend our deepest gratitude to the respected editor and the journal's editorial team for their support and professional handling of our manuscript.

References

[1] A. Bhardwaj, G. V. B. Subrahmanyam, V. Avasthi, and H. Sastry, “Security Algorithms for Cloud Computing,” *Procedia Comput. Sci.*, vol. 85, no. Cms, pp. 535–542, 2016, doi: 10.1016/j.procs.2016.05.215.

[2] B. B. Raj and V. Ceronmani Sharmila, “An survey on DNA based cryptography,” *2018 Int. Conf. Emerg. Trends Innov. Eng. Technol. Res. ICETIETR 2018*, pp. 1–3, 2018, doi: 10.1109/ICETIETR.2018.8529075.

[3] C. Paar and J. Pelzl, *Understanding Cryptography*. 2010. doi: 10.1007/978-3-642-04101-3.

[4] J. T. Isaac and S. Zeadally, “Secure Mobile Payment Systems,” *IT Prof.*, vol. 16, no. June, pp. 36–43, 2014, doi: 10.1109/MITP.2014.40.

[5] M. MONDAL and K. S. RAY, “Review on DNA Cryptography,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 2, no. 1, pp. 44–76, 2023, doi: 10.48550/arXiv.1904.05528.

[6] R. Joshi, M. C. Trivedi, V. Goyal, and D. Bhati, *DNA Sequence in Cryptography: A Study*, vol. 522, no. November. Springer Nature Singapore, 2023. doi: 10.1007/978-981-19-5292-0_53.

[7] B. T. Hammad, A. M. Sagheer, I. T. Ahmed, and N. Jamil, “A comparative review on symmetric and asymmetric dna-based cryptography,” *Bull. Electr. Eng. Informatics*, vol. 9, no. 6, pp. 2484–2491, 2020, doi: 10.11591/eei.v9i6.2470.

[8] S. Pramanik and S. K. Setua, "DNA cryptography," *2012 7th Int. Conf. Electr. Comput. Eng. ICECE 2012*, pp. 551–554, 2012, doi: 10.1109/ICECE.2012.6471609.

[9] T. Mandge and V. Choudhary, "A DNA encryption technique based on matrix manipulation and secure key generation scheme," *2013 Int. Conf. Inf. Commun. Embed. Syst. ICICES 2013*, pp. 47–52, 2013, doi: 10.1109/ICICES.2013.6508181.

[10] T. Anwar, A. Kumar, and S. Paul, "DNA Cryptography Based on Symmetric Key Exchange," *Int. J. Eng. Technol.*, vol. 7, no. 3, pp. 938–950, 2015, https://www.enggjournals.com/ijet/docs/IJET15-07-03-321.pdf

[11] V. Kolate and R. Joshi, "An Information Security Using DNA Cryptography along with AES Algorithm," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 1, pp. 183–192, 2021, doi: 10.17762/turcomat.v12i1S.1607.

[12] Madhusudhan R and Shashidhara R, "A novel DNA based password authentication system for global roaming in resource-limited mobile environments," *Multimed. Tools Appl.*, vol. 79, no. 3–4, pp. 2185–2212, 2020, doi: 10.1007/s11042-019-08349-8.

Declaration of Competing Interest: Authors do not have conflict of interest. The content of the paper is approved by the authors.

Author Contributions:
Fatemeh Alidadi Shamsabadi: Methodology, writing original draft preparation, resources, evaluation; **Shaghayegh Bakhtiari Chehelcheshmeh:** Supervision, methodology, resources, manuscript editing; **Narges Farhadi:** Implementation in the Visual Studio.

E-ISSN: 2980-9231
https://sanad.iau.ir/journal/jce

مهندسی مخابرات جنوب
دانشگاه آزاد اسلامی واحد بوشهر

https://doi.org/10.30495/jce.2024.2003925.1242

مقاله پژوهشی

یک سیستم رمز DNA جدید برای رمزگذاری متون فارسی و تصاویر با تبادل کلید مبتنی بر منحنی فرابیضوی

فاطمه علی‌دادی شمس‌آبادی[۱] | شقایق بختیاری چهل‌چشمه*[۲] | نرگس فرهادی[۳]

[۱]گروه کامپیوتر، واحد شهرکرد، دانشگاه آزاد اسلامی، شهرکرد، ایران.
fa.alidadi@iau.ir

[۲]گروه کامپیوتر، واحد شهرکرد، دانشگاه آزاد اسلامی، شهرکرد، ایران.
bakhtiari@iau.ac.ir

[۳]گروه کامپیوتر، واحد شهرکرد، دانشگاه آزاد اسلامی، شهرکرد، ایران.
n.farhadi562@iau.ac.ir

نویسنده مسئول:
*شقایق بختیاری چهل‌چشمه، استادیار گروه کامپیوتر، واحد شهرکرد، دانشگاه آزاد اسلامی، شهرکرد، ایران.
bakhtiari@iau.ac.ir

موضوع اصلی:
رمزنگاری DNA

تاریخچه مقاله:
تاریخ دریافت: ۲۵ آذر ۱۴۰۲
تاریخ بازنگری: ۸ بهمن ۱۴۰۲
تاریخ پذیرش: ۱۱ بهمن ۱۴۰۲

چکیده:

با ظهور کامپیوترهای کوانتومی، روش‌های رمزنگاری سنتی شکسته خواهند شد و کارایی خود را از دست خواهند داد. خوشبختانه، روش‌های رمزنگاری مبتنی بر DNA در برابر حمله‌های کوانتومی مقاومت بالایی دارند. با این حال، تاکنون از این روش‌ها برای حفظ امنیت متون فارسی استفاده نشده است. از سوی دیگر، روش‌های رمزنگاری DNA به‌منظور اشتراک‌گذاری کلید راز، به یک کانال ارتباطی امن نیاز دارند. از این‌رو در این مقاله، یک سیستم رمز جدید به‌منظور تـأمین محرمـانگی متون فارسی بر اسـاس رمزنگاری DNA ارائه می‌شود. همچنین با توجه به استفاده گسترده تصاویر در رسانه‌های عمومی، قابلیت رمزگذاری تصاویر نیز در این طرح نهادینه شده است. در روش پیشنهادی از سیستم رمزنگاری منحنی فرابیضوی (HECC) برای تولید کلید راز اشتراکی در هر دو طرف ارتباط استفاده می‌شود تا دیگر نیازی به یک کانال امن نباشد. همچنین، از فناوری هیبریدسازی DNA به‌منظور انجام فرآیند رمزگذاری و رمزگشایی پیام استفاده می‌شود تا پیچیدگی زمانی الگوریتم رمزنگاری به حداقل برسد. براساس نتایج حاصل از ارزیابی کارایی، طرح پیشنهادی نسبت به طرح‌های قبلی، سربار محاسباتی را کاهش می‌دهد و مصرف انرژی کم‌تری دارد.

کلید واژه‌ها: امنیت، توافق کلید منحنی فرابیضوی، رمزنگاری DNA، متون ساده فارسی، هیبریدسازی DNA.

تازه‌های تحقیق:

- رمزگذاری متون فارسی با استفاده از سیستم رمز DNA برای اولین‌بار
- ارائه فضای کلید بزرگ به‌منظور مقاومت بیشتر در برابر عوامل خارجی
- به حداقل رساندن پیچیدگی زمانی با استفاده از فناوری هیبریدسازی DNA
- عدم نیاز به یک کانال امن به‌منظور اشتراک‌گذاری کلید راز
- توافق کلید بسیار امن و کارا با استفاده از رمزنگاری مبتنی بر منحنی فرابیضوی

۱- مقدمه

با توسعه سریع فناوری اطلاعات و شبکه‌های کامپیوتری، انتقال داده‌های دیجیتالی (به‌ویژه فایل‌های متنی و تصویری) افزایش یافته است. با این‌حال، نگرانی‌های بسیاری در رابطه با تأمین امنیت مربوط به این داده‌ها و محافظت از آن‌ها در برابر کاربران غیرمجاز وجود دارد. رمزنگاری، علمی است که امنیت داده‌ها و ارتباطات روی شبکه را با استفاده از ریاضیات و منطق فراهم می‌کند. به‌کمک رمزگذاری می‌توان پیام ساده را به پیام غیرقابل فهم تبدیل کرد؛ به‌طوری‌که فقط افرادی که مجاز هستند بتوانند از طریق رمزگشایی به محتوای پیام دست پیدا کنند [۱].

تاکنون، الگوریتم‌ها و تکنیک‌های رمزنگاری زیادی برای انتقال امن داده‌ها ارائه شده است. سیستم‌های رمز کلید متقارن که بسیار سریع و محبوب هستند، از کلید یکسانی به‌منظور رمزگذاری و رمزگشایی اطلاعات استفاده می‌کنند. انتقال کلید راز به‌عنوان نقطه ضعف بزرگ رمزگذاری کلید متقارن شناخته می‌شود. در روش رمزگذاری نامتقارن، مسئله انتقال کلید راز برطرف شده است. در سیستم‌های رمز کلید نامتقارن، از یک جفت کلید عمومی و خصوصی برای این منظور استفاده می‌شود. کلید عمومی با همه شرکت‌کنندگان حاضر در ارتباط به اشتراک گذاشته می‌شود. شرکت‌کنندگان به‌وسیله کلید عمومی گیرنده می‌توانند عملیات رمزگذاری را انجام دهند. ولی، کلید خصوصی در نزد گیرنده مورد نظر مخفی می‌ماند. گیرنده پیام، از کلید خصوصی خود برای عملیات رمزگشایی استفاده می‌کند. به‌منظور انتقال امن حجم زیادی از داده‌ها، کاربران به‌طور معمول از تکنیک‌های رمزگذاری کلید نامتقارن برای برقراری ارتباط استفاده می‌کنند. سپس، کلید راز متقارن را به اشتراک می‌گذارند و از رمزنگاری متقارن استفاده می‌کنند. بدین ترتیب، از مزایای هر دو روش رمزنگاری تا حدی بهره می‌برند [۲].

پس از اختراع محاسبات DNA[1] و محاسبات کوانتومی، برخی از الگوریتم‌های رمزنگاری سنتی (مانند DES[2]، AES[3] و غیره) شکسته شده است. از سوی دیگر در فناوری مدرن، به توان محاسباتی و ظرفیت ذخیره‌سازی بالایی برای حجم زیادی از اطلاعات نیاز است [۳]. از این‌رو، زمان آن فرا رسیده است که از جایگزین مناسبی برای سیستم‌های رمز سنتی استفاده شود.

به‌تازگی، سیستم‌های رمز بیولوژیکی[4] به‌طور گسترده‌ای برای کاربردهای مختلف مورد استفاده قرار گرفته‌اند که از میان آن‌ها، رمزنگاری DNA امن‌ترین فناوری برای ارتباطات است. رمزنگاری DNA، آخرین فناوری رمزنگاری است که از فرآیند طبیعی شکل‌گیری DNA به‌منظور دستیابی به یک سیستم امن استفاده می‌کند [۴]. مولکول DNA که قابلیت ذخیره‌سازی، پردازش و انتقال اطلاعات را دارد، ایده رمزنگاری DNA را الهام بخشیده است. یکی از مزایای محاسبات DNA، موازی‌بودن گسترده مولکول‌های DNA است. نتایج یک آزمایش در آزمایشگاه نشان داده است که حدود ۱۰۱۸ پردازنده که به‌طور موازی کار می‌کنند را می‌توان به‌راحتی کنترل کرد. به‌دلیل این موازی‌سازی گسترده، تابع درگاهی[5] که راز اصلی امنیت در بیشتر سیستم‌های رمز سنتی است، می‌تواند در زمان چند جمله‌ای حل شود. از آن‌جایی که جنبه ریاضی رمزنگاری با شیمی DNA در زمینه رمزنگاری DNA جایگزین می‌شود؛ پس این تکنیک با روش‌های مرسوم و همچنین، محاسبات کوانتومی عملاً غیرقابل هک است [۲]. بنابراین، رمزنگاری DNA را می‌توان جزو سیستم‌های رمز پس‌کوانتومی به‌شمار آورد؛ چراکه در مقابل انواع حمله‌ها به‌وسیله کامپیوتر (الگوریتم) کوانتومی مقاوم است [۵].

در ضمن، مولکول‌های DNA ظرفیت ذخیره‌سازی زیادی دارند. هر گرم مولکول DNA از ۱۰۲۱ باز DNA تشکیل شده است که تقریباً برابر با ۱۰۸ ترابایت است. هر گرم DNA، قابلیت ذخیره‌سازی همان مقدار از اطلاعاتی را دارد که ممکن است در یک تریلیون دیسک فشرده (CD)[6] قرار گیرد. بنابراین، می‌توان نتیجه گرفت که چند گرم DNA می‌تواند تمامی داده‌های ذخیره‌شده در جهان را نگه دارد. این موضوع، در کاهش پیچیدگی فضایی در رمزنگاری DNA تأثیر به‌سزایی دارد [۶, ۷]. رمزنگاری DNA، موازی‌سازی گسترده و ظرفیت ذخیره‌سازی بالای مولکول‌های DNA را با روش‌های سنتی رمزنگاری (متقارن و نامتقارن) ترکیب می‌کند [۲]. با این حال، نیاز به کانال امن یا راه‌حلی امن و کارا برای توافق کلید، از چالش‌های موجود در سیستم‌های رمز DNA به‌شمار می‌رود.

---

[1] Deoxyribo Nucleic Acid
[2] Data Encryption System
[3] Advanced Encryption System
[4] Biological
[5] Trapdoor
[6] Compact Disc

منحنی فرابیضوی، یکی از سیستم‌های رمز برای توافق کلید است که سرعت بسیار بالایی داشته و در گره‌های سیار با منابع محدود، برای دستیابی به امنیت بهتر و بهره‌وری انرژی بیشتر مورد استفاده قرار می‌گیرد. همچنین، در مقایسه با سایر روش‌های رمزنگاری (نظیر RSA[۱]، DES، ECC و غیره) کارایی عملیاتی بیشتر و سربار محاسباتی کم‌تری دارد. HECC[۲] می‌تواند به‌راحتی برای کاربردهای سخت‌افزاری و نرم‌افزاری پیاده‌سازی شود [۳].

هدف اصلی این پژوهش، ارائه یک طرح جدید به‌منظور رمزنگاری ترکیبی متون ساده فارسی و تصاویر با استفاده از سیستم رمز HECC و رمزنگاری DNA است؛ به‌طوری‌که امنیت بالای اطلاعات تضمین شود و سربار محاسباتی کاهش پیدا کند. بخش‌های بعدی این پژوهش به‌صورت زیر سازماندهی شده است: در بخش ۲ آثار قبلی در این زمینه بررسی می‌شود. در بخش ۳ به معرفی رمزنگاری DNA و فناوری‌های قابل استفاده در آن‌ها پرداخته می‌شود. در بخش ۴ روش پیشنهادی ارائه می‌شود و در بخش ۵ امنیت و کارایی سیستم رمز پیشنهادی مورد ارزیابی قرار می‌گیرد. بخش ۶ به بحث و نتیجه‌گیری نهایی اختصاص می‌یابد.

۲- کارهای قبلی

بسیاری از محققان، سیستم‌های رمز پیشرفته مختلفی را با استفاده از محاسبات DNA برای محافظت از فایل‌های محرمانه کاربران پیشنهاد کرده‌اند. در محاسبات DNA، عملیات منطقی و محاسباتی با استفاده از ویژگی‌های مولکولی DNA و به‌وسیله تراشه‌های زیستی به‌جای تراشه‌های کربنی یا سیلیکونی انجام می‌شود [۸]. خود محاسبات DNA در سال ۱۹۹۴ توسط ادلمن[۳] مطرح شد. ادلمن که از ارائه‌دهندگان الگوریتم RSA نیز به‌شمار می‌رود، توانست مسئله مسیر همیلتونی (HPP)[۴] را با استفاده از عملیات DNA حل کند [۹]. محاسبات DNA در زمینه‌های رمزنگاری و پنهان‌نگاری[۵] بیشترین کاربرد را دارد.

توجه داشته باشید که پنهان‌نگاری با رمزنگاری فرق دارد. به‌منظور درک بهتر این موضوع، یک زندانی را در نظر بگیرید که می‌خواهد برای به‌اشتراک‌گذاشتن نقشه فرار از زندان، با شخص دیگری ارتباط برقرار کند. در این‌جا، وجود نگهبانان زندان یا نیروهای پلیس، چالش برانگیز است. در چنین شرایطی باید هرگونه ارتباطی پنهان بماند. بدین‌ترتیب، پنهان‌نگاری برای موقعیت‌هایی که ارتباط ناامن است، یک روش مفید و کارا به‌حساب می‌آید. از طرف دیگر، فرض کنید که یکی از همان نیروهای پلیس می‌خواهد پیامی را با رئیس ارشد خود به‌اشتراک بگذارد. طبیعی است که نیروهای پلیس با رئیس ارشد خودشان ارتباط دارند. بنابراین، خود ارتباط هیچ‌گونه سوءظنی را ایجاد نمی‌کند. با این‌حال، از آن‌جایی که محتوای پیام محرمانه است، ممکن است که برای جلوگیری از شنود جاسوسان یا مهاجمان، از رمزنگاری استفاده شود. با توجه به آن‌چه که گفته شد، هدف از رمزنگاری، تأمین امنیت محتویات پیامی است که قابل‌مشاهده است. در حالی که هدف از پنهان‌نگاری، مخفی‌کردن هرگونه نشانه‌ای از وجود پیام است که برای این منظور، اطلاعات درون یک حامل جاسازی می‌شود تا قابل دیده‌شدن نباشند. ترکیب پنهان‌نگاری با رمزنگاری برای انتقال اطلاعات، بسیار مطلوب است [۱۰]. در ادامه، به بررسی جدیدترین آثار موجود در این زمینه پرداخته می‌شود.

مونیکا و آپادیایا[۶] [۱۱] در سال ۲۰۱۵ یک طرح رمزنگاری DNA با لایه سوکت امن (SSL)[۷] را برای شبکه‌های حسگر بی‌سیم پیشنهاد کردند. به‌منظور تأمین امنیت در این طرح، از جفت کلید (عمومی و خصوصی) استفاده شده است. جفت کلید مورد استفاده برای رمزگذاری و رمزگشایی به‌وسیله الگوریتم RSA تولید می‌شوند. کلیدها و گواهی‌نامه‌های دیجیتالی با توجه به ویژگی‌های شبکه‌های حسگر (مانند فقدان زیرساخت مطمئن و محدودیت‌های منابع) قبل از استقرار در هر محیطی بین تمام گره‌های حسگر توزیع می‌شود. پس از استقرار گره‌های حسگر، هر گره‌ای دارای یک جفت کلید عمومی و خصوصی و خود گواهی‌نامه دیجیتالی است. تبادل کلیدهای عمومی و گواهی‌نامه‌های دیجیتالی بین گره‌های حسگر از طریق کانال امن SSL انجام می‌شود. در این طرح، متن ساده به کد ASCII[۸] تبدیل می‌شود. سپس، با کلید عمومی گیرنده رمزگذاری‌شده و به‌شکل

[1] Rivest Shamir Adleman
[2] Hyper Elliptic Curve Cryptography
[3] Adleman
[4] Hamiltonian Path Problem
[5] Steganography
[6] Monika and Upadhyaya
[7] Secure Socket Layer
[8] American Standard Code for Information Interchange

دودویی تبدیل می‌شود. درنهایت، این رشته دودویی با استفاده از فناوری کدگذاری دیجیتالی DNA به‌صورت توالی DNA درمی‌آید. فرآیند رمزگشایی، برعکس همین فرآیند با استفاده از کلید خصوصی گیرنده انجام می‌شود. از آن‌جایی که توزیع اطلاعات کلیدها به‌وسیله SSL و قبل از استقرار گره‌های حسگر در محیط انجام می‌شود، مصرف انرژی در گره‌های حسگر کاهش می‌یابد. با این حال، مسائل مربوط به رمزنگاری کلید عمومی به‌ویژه هزینه هنگفت مدیریت گواهی‌نامه‌ها و مسئله سپردن کلید مطرح می‌شود.

بارمن و سحا[1] [۱۲] در سال ۲۰۱۵ یک طرح رمزگذاری منحنی بیضوی کدگذاری شده با DNA را ارائه کردند. این رویکرد مبتنی بر DNA-ECC دارای دو سطح امنیتی به‌طور هم‌زمان است؛ سطح اول، کدگذاری مبتنی بر توالی DNA ناشناخته است و سطح دوم، سیستم رمزگذاری و رمزگشایی مبتنی بر ECC است. نحوه عملکرد این الگوریتم بدین‌صورت است که در ابتدا، متن ساده به مقدار ASCII معادل خود تبدیل می‌شود. مقادیر ASCII به اعداد دودویی تبدیل می‌شوند. به‌طور محرمانه، یک توالی DNA از میان توالی‌های در دسترس عموم انتخاب می‌شود که آن را هم فرستنده و هم گیرنده می‌دانند. این بازهای نوکلئوتیدی با استفاده از کدگذاری دیجیتالی DNA به‌صورت دودویی تبدیل می‌شوند. توالی DNA دودویی‌شده به چند قسمت تقسیم می‌شود؛ به‌طوری‌که هر بخش، شامل تعداد بیت‌هایی است که به‌طور تصادفی انتخاب شده‌اند. هر بیت از متن ساده دودویی در ابتدای بخش‌های دودویی از توالی نوکلئوتیدی درج می‌شود. بخش‌ها دوباره به‌هم متصل شده و به حروف نوکلئوتیدی تبدیل می‌شوند. توالیهای جدید به مبنای ۱۰ برده می‌شوند. آن‌گاه با استفاده از روش کوبلیتز[2]، اعداد در مبنای ۱۰ به نقطه منحنی بیضوی تبدیل می‌شوند تا نقطه متن رمزی به‌دست آید. این نقطه برای گیرنده ارسال می‌شود. گیرنده در فرآیندی معکوس می‌تواند پیام متن ساده را به‌دست آورد. بدیهی است که عملیات رمزگذاری و رمزگشایی در ECC به‌کمک کلیدهای تولید شده در آن انجام می‌شود. از مزایای این طرح می‌توان به سربار محاسباتی پایین و مصرف بهینه حافظه اشاره کرد. با این حال، این طرح به‌دلیل اندازه کلید کوچک، در برابر حمله جستجوی فراگیر[3] چندان امن نیست. علاوه‌براین، در مرجع [۳] اثبات شده است که کارایی ECC نسبت به رمزنگاری منحنی فرابیضوی کم‌تر است.

انور[4] و همکاران [۶] در سال ۲۰۱۶ به‌منظور محاسبه توالی DNA مربوط به پد یک‌بار مصرف (OTP)[5] از روش تبادل کلید متقارن بهره بردند. این پد یک‌بار مصرف برای رمزنگاری، فقط یک‌بار مورد استفاده قرار می‌گیرد. فرآیند رمزگذاری در این الگوریتم بدین صورت است که ابتدا، متن ساده به کد ASCII تبدیل می‌شود. این کد ASCII به‌صورت یک رشته دودویی در می‌آید و از سمت چپ به راست در یک ماتریس قرار می‌گیرد. سپس، ستون‌ها به‌سمت راست جابه‌جا می‌شوند تا تصویر آینه‌ای از آن ماتریس به‌دست آید. علاوه‌براین، توالی کلید DNA به‌شکل دودویی درمی‌آید و به‌صورت ستونی در یک ماتریس درج می‌شود. عملیات XOR بین ماتریس‌های به‌دست‌آمده انجام می‌شود تا ماتریس جدیدی به‌وجود آید. عناصر ماتریس حاصل به‌صورت ستونی در یک ردیف قرار می‌گیرند تا یک رشته دودویی به‌وجود آید. موقعیت بیت‌های ۱ از این رشته دودویی، در توالی کلید DNA مشخص می‌شود و توالی DNA مربوطه از روی توالی کلید DNA به‌دست می‌آید. این توالی DNA محاسبه شده و آخرین بیت از آن حذف می‌شود. این کار برای تمامی بیت‌های ۱ تکرار می‌شود. درنهایت، تمام آن‌ها کنار هم قرار می‌گیرند تا متن رمزی تولید شود. فرستنده، متن رمزی را در قالب بسته‌هایی برای گیرنده می‌فرستد. گیرنده نیز از روشی معکوس برای رمزگشایی متن رمزی استفاده می‌کند. همان‌طور که پیداست، روند انجام این الگوریتم بسیار طولانی است.

پوجاری[6] و همکاران [۱۳] در سال ۲۰۱۸ از مفاهیم الگوریتم ژنتیک و رمزنگاری DNA به‌منظور رمزگذاری تصاویر استفاده کردند. این طرح مبتنی بر رمزنگاری متقارن است که از دو مرحله درهم‌سازی و جایگزینی تشکیل شده است. در مرحله اول، مکان پیکسل‌ها با استفاده از الگوریتم ژنتیک تغییر می‌کند تا همبستگی میان پیکسل‌های مجاور کاهش یابد. در مرحله جایگزینی، مقدار پیکسل‌ها با مقدار دیگری جایگزین می‌شود؛ برای این کار، ابتدا مقدار پیکسل‌ها به رشته‌های دودویی تبدیل شده و سپس، مقدار پیکسل‌ها با نتیجه عملیات XOR بین این رشته‌های دودویی و زیررشته‌های DNA جایگزین می‌شود.

---

[1] Barman and Saha
[2] Koblitz
[3] Brute Force
[4] Anwar
[5] One-Time Pad
[6] Pujari

زیررشته‌های DNA نیز از یک رشته تصادفی DNA به‌دست می‌آیند. هر زیررشته DNA به‌عنوان یک کلید راز اشتراکی برای رمزنگاری تصویر مورد استفاده قرار می‌گیرد. با این وجود، این طرح به یک کانال امن برای تبادل کلید راز نیاز دارد.

ویکاس[1] و همکاران [۱۴] در سال ۲۰۱۸ الگوریتمی را برای تولید کلید براساس بخشی از توالی DNA و بخشی از عدد خیلی بزرگ PI پیشنهاد دادند تا براساس آن کلید، عملیات رمزگذاری و رمزگشایی انجام شود. جالب این‌جا است که در حدود ۳/۳ میلیارد ترکیب مختلف از رشته‌های DNA وجود دارد. علاوه‌براین، دانشمندان میلیاردها رقم PI را محاسبه کرده‌اند. مقدار PI با ۳/۱۴۱۵۹۲۶۵۳۵ شروع می‌شود؛ ولی از آن‌جایی‌که هیچ الگوی خاصی در جانشینی ارقام آن وجود ندارد، می‌توان محاسبه ارقام بعدی را برای هزاران سال ادامه داد! بنابراین، فضای کلید بزرگی وجود دارد که قابلیت تصادفی‌بودن کلید را نیز فراهم می‌کند. برای تولید کلید مشترک، فرستنده یک موقعیت از رشته DNA (مثلاً نوکلئوتید ۲۰ تا ۲۳) و یک موقعیت از عدد PI (مثلاً رقم ۲۰ تا ۲۳) را انتخاب کرده و آن‌ها را برای گیرنده ارسال می‌کند. گیرنده، این رشته DNA را مطابق با فناوری کدگذاری دیجیتالی DNA به‌صورت دودویی تبدیل کرده و مکمل آن‌را به‌دست می‌آورد. در نهایت، روی مکمل کد دودویی با عدد PI که به‌شکل دودویی درآمده، عملیات XOR را انجام می‌دهد تا به کلید دست یابد. نتایج حاصل از ارزیابی کارایی نشان‌داده‌است که این طرح در مقایسه با سیستم‌های رمز AES و DES سرعت بیشتری دارد. البته در این الگوریتم، انتقال اطلاعات مورد نیاز برای تولید کلید (رشته DNA و عدد PI انتخابی) چالش‌برانگیز است؛ چراکه اگر این اطلاعات در اختیار افراد غیرمجاز قرار گیرد، به‌راحتی می‌توانند کلید راز را به‌دست آورده و این سیستم رمز را بشکنند. متأسفانه در این سیستم رمز به‌منظور انتقال چنین اطلاعات حساسی، هیچ مکانیزم امنیتی ارائه نشده است.

شرعیه[2] و همکاران [۱۵] در سال ۲۰۱۹ یک الگوریتم رمز چند الفبایی پیشرفته (EPCA)[3] را با استفاده از رمزنگاری DNA ارائه نمودند. رمز چند الفبایی، نوعی رمز چند جایگزینی است؛ به‌طوری‌که در آن، هر حرف از متن ساده با شکل‌های مختلفی جایگزین می‌شود. محدودیت اصلی در رمزهای چند الفبایی این است که فقط از حروف انگلیسی استفاده می‌کنند. این موضوع باعث می‌شود که با تحلیل فراوانی، بتوان حروف متن رمزی را حدس زد. به‌منظور برطرف‌کردن این مسئله، EPCA از کاراکترها و علائم دیگری نیز پشتیبانی می‌کند. در این الگوریتم برای تولید کلید، دو عدد تصادفی بین ۶۵ و ۱۲۲ انتخاب می‌شود. این اعداد تصادفی به‌عنوان کلیدهای راز به یونیکد[4] دودویی تبدیل می‌شوند. به‌منظور مبهم‌شدن هرچه بیشتر کلید راز در فرآیندهای رمزگذاری و رمزگشایی، از محاسبات و عملگر بیتی روی این دو عدد تصادفی استفاده‌شده‌است. فرآیند رمزگذاری در این طرح، از سه کلید راز استفاده می‌کند؛ یکی برای رمزگذاری متن ساده و دو کلید دیگر برای رمزگذاری و رمزگشایی کلیدهایی که به شروع و انتهای متن رمزی اضافه می‌شوند. نحوه عملکرد این سیستم رمز بدین صورت است که ابتدا متن ساده به داده‌های دودویی تبدیل می‌شود. داده‌های دودویی حاصل از متن ساده با استفاده از فناوری کدگذاری دیجیتالی DNA به یک رشته DNA تبدیل می‌شوند. اکنون، EPCA با استفاده از کلید راز روی رشته DNA مذکور اعمال می‌شود. خروجی فرآیند رمزگذاری، متن رمزی با کلیدهای متقارن رمزشده است. فرآیند رمزگشایی در این طرح، طبق روندی معکوس صورت می‌گیرد. مشخص است که این طرح، سربار محاسباتی زیادی دارد.

پاتنالا و کومار[5] [۱۶] در سال ۲۰۱۹ یک الگوریتم رمزگذاری براساس کدون‌های DNA را پیشنهاد کردند. هر سه نوکلئوتید پشت سرهم، یک کدون را تشکیل می‌دهد. مجموع کامل کدون‌ها نیز کد ژنتیکی را می‌سازد. در این الگوریتم با انتخاب هر سه نوکلئوتید از رشته DNA می‌توان ۶۴ ترکیب از کدون‌ها را به‌وجود آورد. هدف این طرح، رمزگذاری یک متن‌ساده شامل حروف بزرگ یا کوچک انگلیسی، اعداد صفر تا نه، فاصله و نقطه است که در مجموع ۶۴ کاراکتر می‌شود. این طرح از روش جایگزینی برمبنای یک جدول جستجو استفاده می‌کند. جدول جستجو، دربرگیرنده کدون‌های DNA و کاراکترهای متناظر با آن‌ها است که به‌طور تصادفی مرتب شده‌اند. متن ساده به‌منظور رمزگذاری، به کد ASCII تبدیل می‌شود. کد ASCII به‌شکل اعداد دودویی درمی‌آید. هر کد دودویی به‌صورت دو بیتی تقسیم می‌شود. هر دو بیت از کد دودویی به‌کمک فناوری کدگذاری دیجیتالی DNA

[1] Vikas
[2] Namasudra and Deka
[3] Enhanced Polyalphabetic Cipher Algorithm
[4] Unicode
[5] Patnala and Kumar

به توالی DNA تبدیل می‌شود. اکنون، از هر سه نوکلئوتید این توالی DNA برای ایجاد یک کدون استفاده می‌شود. پس از آن، هر کدون با کاراکتر متناظر خود از جدول جستجو جایگزین می‌شود. کاراکترهای به‌دست‌آمده، به کد ASCII تبدیل می‌شوند. کد ASCII به رشته دودویی تبدیل شده و دوباره، به‌صورت دو بیتی تقسیم می‌شود. در نهایت، هر دو بیت از رشته دودویی با استفاده از کدگذاری دیجیتالی DNA به‌صورت توالی DNA در می‌آید که همان، متن رمزی است. فرآیند رمزگشایی نیز به‌صورت برعکس انجام می‌شود. نتایج حاصل از شبیه‌سازی این طرح به‌وسیله برنامه‌نویسی net. حاکی از آن است که با افزایش طول متن ساده، زمان رمزنگاری به‌طور چشم‌گیری افزایش می‌یابد. متأسفانه، این الگوریتم به‌دلیل انجام چندین دور با عملیات‌های تکراری بسیار زمان‌بر است.

صالح و همکاران [۱۷] در سال ۲۰۲۰ الگوریتمی را به‌منظور پنهان‌سازی متن ساده در یک فایل تصویری ارائه کردند. در این‌جا فرض شده است که اندازه تصویر حامل پیام، (۲۵۶,۲۵۶) است. طول خود پیام برای پنهان‌نگاری نیز یک کیلوبایت در نظر گرفته‌شده‌است. ولی پیش از پنهان‌سازی، متن ساده با استفاده از روش جانشینی الفبا رمز می‌شود. این متن رمزی به‌کمک الگوریتم ژنتیک تنظیم می‌شود. آن‌گاه، فناوری کدگذاری دیجیتالی DNA برای تبدیل‌کردن متن رمزی به‌منظور جاسازی در تصویر با استفاده از بیت کم‌ارزش به‌کار می‌رود. اکنون، تصویر حامل پیام از طریق یک کانال ارتباطی برای گیرنده فرستاده می‌شود. گیرنده، با استفاده از کلید راز اشتراکی می‌تواند پیام جاسازی‌شده در تصویر حامل را استخراج کند. سپس، به روشی معکوس به متن ساده دست یابد. اگرچه این الگوریتم پس از پنهان‌نگاری، تأثیر چندانی روی کیفیت تصویر ندارد؛ ولی، راه‌کاری را به‌منظور توافق کلید راز اشتراکی ارائه نکرده است.

وینوتا و جوز[1] [۱۸] در سال ۲۰۲۰ برای کدگذاری داده‌های یک تصویر از فناوری واکنش زنجیره‌ای پلیمراز (PCR)[2] استفاده کردند. برای این منظور، داده‌های کدشده بین کلیدهای پرایمر قرار می‌گیرد. کلیدهای پرایمر به‌کمک فناوری OTP توسط مولد توالی شبه‌تصادفی تولید می‌شوند. همچنین، توالی DNA و کلیدهای رمزگذاری با استفاده از کدون DNA به‌صورت بازهای چهارتایی تولید می‌شوند؛ به‌طوری‌که به‌صورت ریز نقاط یا هر فرآیند ارتباطی دیگری بین فرستنده و گیرنده به اشتراک گذاشته می‌شوند. مراحل رمزگذاری در این الگوریتم از این قرار است که یک تصویر به‌عنوان ورودی گرفته می‌شود. تصویر با استفاده از متلب[3] به کد دودویی تبدیل می‌شود. هر مقدار پیکسلی از تصویر به‌صورت کد دودویی ۸ بیتی درمی‌آید. داده‌های دودویی با استفاده از فناوری کدگذاری دیجیتالی DNA به کد DNA تبدیل می‌شوند. کد DNA با استفاده از کلید رمزگذاری، به متن رمزی تبدیل می‌شود. کلید پرایمر به‌شکل توالی DNA به متن رمزی تبدیل می‌شود و پیام کدگذاری‌شده در بین این کلیدهای پرایمر قرار می‌گیرد. متن رمزی با استفاده از متلب به تصویر رمزگذاری‌شده تبدیل می‌شود. این طرح، دامنه کلید را افزایش داده است تا احتمال حدس‌زدن کلید توسط مهاجم کاهش پیدا کند. ولی، این سیستم به‌علت استفاده از فناوری‌های PCR و OTP وقت‌گیر است و تنها برای پیام‌های کوتاه قابل اجرا است.

عالم و سینگ[4] [۱۹] در سال ۲۰۲۱ چالش‌هایی که مانع از پذیرش سیستم‌های ابری توسط افراد و سازمان‌ها می‌شوند را بررسی نمودند. بیشتر این نگرانی‌ها به حفظ حریم خصوصی و امنیت اطلاعات مرتبط با ابر برمی‌گردد. از آن‌جایی که داده‌ها در اختیار سوم شخص قرار می‌گیرند، صاحبان داده روی داده‌های برون‌سپاری‌شده خود کنترلی ندارند. همچنین، صاحبان داده باید ریسک زیادی را بپذیرند تا به ارائه‌دهنده سرویس ابری اعتماد کنند. به‌همین‌منظور در این مقاله، یک سیستم رمز DNA چندلایه‌ای برای فرآیند رمزگذاری و رمزگشایی داده‌های برون‌سپاری‌شده به ابر طراحی و پیاده‌سازی است؛ به‌طوری که امکان رمزنگاری را به‌طور مستقیم به کاربران مجاز در دستگاه مربوط به خودشان می‌دهد. برای انجام فرآیند رمزگذاری، کاراکترهای پیام اصلی به توالی DNA تبدیل می‌شود. دوباره، این توالی DNA به بازهای نوکلئوتیدی دیگری تبدیل می‌شود. سپس، این بازهای نوکلئوتیدی به‌شکل کد دودویی درمی‌آید. آن‌گاه، یک عدد تصادفی تولید می‌شود و با این کد دودویی XOR می‌شود. کد دودویی حاصل از عملیات XOR، به یک عدد در مبنای ۱۰ و پس آن به کاراکتر ASCII معادل تبدیل می‌شود. درنهایت، این کاراکتر ASCII برای سرویس‌دهنده ابری ارسال می‌شود. گیرنده پس از دانلود فایل از روی سرویس‌دهنده ابری، همین مراحل را به‌طور معکوس انجام

[1] Vinotha and Jose
[2] Polymerase Chain Reaction
[3] MATLAB
[4] Alam and Singh

می‌دهد تا پیام رمزگشایی شود. بنابراین، کاربر می‌تواند جامعیت داده‌ها را بدون نیاز به سوم شخص بررسی کند. علاوه‌براین، درستی این الگوریتم با پیاده‌سازی در پایتون اثبات شده است. محرمانگی داده‌ها در ابر تضمین می‌شود؛ چراکه این داده‌ها به‌طور رمز شده در اختیار سرویس‌دهنده ابری قرار می‌گیرد. با این حال، مسئله توزیع کلید در این طرح حل نشده است.

پراسانا[1] و همکاران [۲۰] در سال ۲۰۲۲ رویکردی را به‌منظور ذخیره‌سازی داده‌های رمزگذاری‌شده به‌وسیله محاسبات DNA در ابرهای چندگانه ارائه کردند. در این طرح، داده‌های حیاتی کاربر به قطعه‌های کوچک‌تری تقسیم می‌شوند. سپس، این قطعه‌های داده که به‌وسیله محاسبات DNA رمزگذاری شده‌اند، بین سرویس‌دهندگان ابری واجد شرایط توزیع می‌شوند. مراحل انجام این الگوریتم به دو فاز جاسازی و استخراج داده‌ها تقسیم می‌شود. در فاز اول، داده‌های دودویی کاربر با استفاده از فناوری کدگذاری DNA به نوکلئوتیدهای DNA تبدیل می‌شود. پس از آن، قانون جفت‌سازی باز اعمال می‌شود تا شکل جدیدی از نوکلئوتیدهای DNA تولید شود. اکنون، متن رمزی با پیداکردن اندیس نوکلئوتیدها در توالی مرجع DNA به‌دست می‌آید. بدین‌ترتیب، مرحله جاسازی‌کردن داده‌ها به پایان می‌رسد و اطلاعات کاربر به‌صورت رمزشده برای ابر ارسال می‌شود. به‌منظور استخراج این داده‌ها، مراحل جاسازی‌کردن داده‌ها به‌شکل معکوس انجام می‌شود. این تکنیک، فضای ذخیره‌سازی امن‌تری را برای کاربر ابری از طریق تقسیم‌کردن داده‌های حیاتی کاربر به چند قطعه، اعمال رمزنگاری قوی مبتنی‌بر DNA برای هر قطعه از داده‌ها و بارگذاری آن‌ها در ابرهای مختلف فراهم کرده‌است. با این وجود، سربار ارتباطی در این طرح نسبت به سایر طرح‌ها بسیار زیاد است.

الوطار[2] [۲۱] در سال ۲۰۲۳ از داده‌های DNA موجود در بانک ژنتیکی[3] به‌منظور پنهان‌نگاری استفاده کرد. در این معماری، کلید عمومی گیرنده برابر با موقعیت قطعه DNA انتخاب‌شده در بانک ژنتیکی است. در حالی‌که کلید خصوصی او، چند قطعه DNA متناظر با موقعیت تعیین‌شده در بانک ژنتیکی است. این بدان معناست که مقدار قطعه DNA انتخاب‌شده، دارای تعداد معینی از بازهای DNA و طول مشخصی است. گیرنده می‌تواند کلید را با هر روشی تولید یا استفاده کند. فرستنده، موقعیت قطعه را به‌عنوان کلید عمومی گیرنده می‌داند؛ درحالی‌که گیرنده، مقدار و طول آن‌را می‌داند. پنهان‌سازی و استخراج در این طرح بدین صورت است که ابتدا، پیامی توسط فرستنده برای پنهان‌سازی آماده می‌شود. سپس، فایلی برای جاسازی پیام در نظر گرفته می‌شود. پیام، در فایل پوششی با استفاده از کلید عمومی گیرنده جاسازی می‌شود. سپس، پیام مخفی‌شده در فایل پوششی برای گیرنده ارسال می‌شود. گیرنده، از کلید عمومی برای یافتن موقعیت قطعه در بانک ژنتیکی بهره می‌برد و از مقدار قطعه، برای به‌دست‌آوردن کلید خصوصی خود به‌منظور استخراج پیام مخفی‌شده استفاده می‌کند. این کار، بسیار زمان‌بر و پرهزینه است؛ چراکه اصل پنهان‌نگاری در آن، مبتنی‌بر سیستم رمز کلید عمومی است.

بر اساس بررسی‌ها، تاکنون هیچ طرحی برای رمزگذاری متون فارسی با استفاده از رمزنگاری DNA ارائه نشده است. علاوه‌براین، طرح‌های موجود نتوانسته‌اند امنیت را به‌طور ترکیبی هم برای تصاویر و هم برای متون فراهم کنند. در بیشتر طرح‌های بررسی شده، یا هیچ مکانیزم امنیتی برای توزیع کلید ارائه نشده است؛ یا مکانیزم امنیتی که برای توزیع کلید ارائه شده، با سربار محاسباتی زیاد و سرعت اجرای پایین همراه بوده است. نقطه ضعف بزرگ دیگر در این الگوریتم‌ها، فضای کوچک کلید است که مانعی برای تضمین امنیت اطلاعات به‌شمار می‌رود.

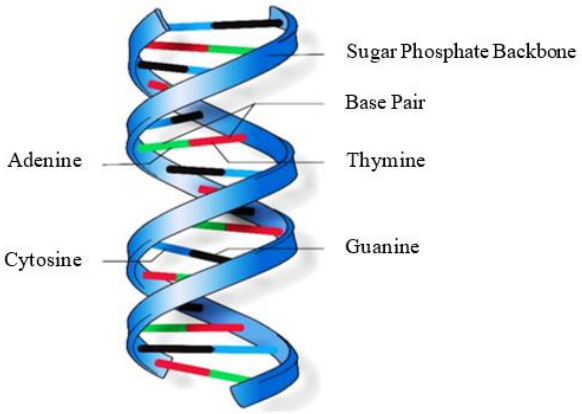
۳- سیستم رمز DNA

رمزنگاری DNA یک مفهوم ترکیبی است که از دو حوزه مختلف، یعنی زیست‌شناسی و علوم کامپیوتر پدید آمده است [۲۲]. در این روش، از DNA به‌عنوان حامل اطلاعات و از فناوری بیولوژیکی مدرن به‌عنوان ابزار پیاده‌سازی استفاده می‌شود. DNA مولکولی است که در همه موجودات زنده یافت می‌شود. وظیفه اصلی آن، انتقال اطلاعات ژنتیکی مورد نیاز برای رشد، نمو و عملکرد از نسلی به نسل دیگر است. اطلاعات در DNA به‌صورت کدی از چهار باز شیمیایی به نام‌های ادنین (A)[4]، گوانین (G)[5]،

[1] Prasanna
[2] AL-Wattar
[3] GenBank
[4] Adenine
[5] Cytosine

سیتوزین (C)[1] و تیمین (T)[2] ذخیره می‌شود. به‌طور کلی، همیشه A با T و C با G پیوندهای هیدروژنی تشکیل می‌دهند. همچنین، هر باز به دو مولکول دیگر به‌نام مولکول‌های قند و فسفات پیوند می‌خورد. یک باز همراه با مولکول‌های قند و فسفات به‌عنوان نوکلئوتید شناخته می‌شود. در رمزنگاری DNA از نحوه قرارگیری همین بازهای نوکلئوتیدی برای رمزکردن اطلاعات استفاده می‌شود [۲۳-۲۵].

همان‌طور که در شکل ۱ نشان داده شده است، نوکلئوتیدها در دو رشته بلند قرار گرفته‌اند که ساختار مارپیچ‌گونه DNA را تشکیل می‌دهند. این ساختار تا حدی شبیه به نردبانی است که میله‌های جانبی عمودی در آن، مولکول‌های قند و فسفات هستند و جفت پایه، پله‌های نردبان را تشکیل می‌دهند [۲۴]. DNA تنها با ترکیبی از این چهار حرف A، C، T و G حجم بسیار پیچیده و زیادی از اطلاعات را برای یک موجود زنده ذخیره می‌کند [۲۶]. مفهوم موازی‌سازی گسترده و تراکم زیاد اطلاعات که به‌طور ذاتی در مولکول‌های DNA وجود دارد، برای اهداف رمزنگاری مورد بهره‌برداری قرار می‌گیرد [۲۷].



شکل ۱: ساختار DNA [۲۲]
Figure 1. Structure of DNA [22]

۳-۱- فناوری‌های DNA

همان‌طور که از جدول ۱ برمی‌آید، سیستم‌های رمز مبتنی بر DNA از انواع مختلف فناوری‌ها به‌منظور تضمین امنیت داده‌ها استفاده می‌کنند. در این بخش به بررسی این فناوری‌ها پرداخته می‌شود.

ـ پد یکبار مصرف تصادفی مبتنی بر DNA: در این تکنیک، یک توالی DNA مخفی به‌طور تصادفی انتخاب می‌شود. فناوری OTP برای رمزگذاری از یک کلید تصادفی استفاده می‌کند که این کلید، هر بار تغییر پیدا می‌کند. به همین دلیل، این تکنیک به‌نام پد یک‌بار مصرف نام‌گذاری شده‌است. این موضوع، به افزایش امنیت کمک می‌کند. در این طرح، طول پد یک‌بار مصرف باید برابر با طول متن ساده باشد. یک کتابچه کد تصادفی و مخصوص به‌منظور جایگزینی متن ساده در نظر گرفته می‌شود. طرح DNA-OTP در بیشتر مواقع برای تبدیل متن ساده کوتاه یا بخشی از یک پیام متنی ساده به متن رمزی استفاده می‌شود. متأسفانه، این روش روی پیام‌های بزرگ قابل اجرا نیست؛ چراکه سخت‌افزار فعلی به اندازه کافی برای پد یکبار مصرف مناسب نیست. شایان ذکر است که افزایش اندازه پیام، پیچیدگی نگاشت DNA را شدت می‌بخشد [۲۶,۲۸].

ـ تراشه مبتنی بر DNA: تراشه DNA به‌طور معمول با عنوان ریزآرایه شناخته می‌شود. ریزآرایه، از اسید نوکلئیک[3] تشکیل شده و مدار الکترونیکی آن، از نیمه‌هادی‌ها ساخته شده است. میلیون‌ها میله DNA در مساحتی کمتر از یک اینچ مربع بر روی ماتریسی سیلیکونی قرار گرفته‌اند. از این تراشه‌ها مانند تراشه‌های سیلیکونی می‌توان برای ذخیره‌سازی داده‌ها در قالب توالی‌های DNA استفاده کرد. رمزگذاری و رمزگشایی با استفاده از فرآیندهای بیوشیمیایی انجام می‌شود و می‌توان آن‌را در رمزگذاری پیام‌های متنی ساده و تصاویر به‌کار برد. از آن‌جایی‌که تراشه DNA یک عنصر بیولوژیکی است، ویژگی‌های آن ممکن است به‌دلیل سایر عوامل فیزیکی تغییر کند. این تغییر ناگهانی، تأثیر منفی روی پیام‌های رمزگذاری‌شده خواهد داشت [۲۹,۳۰].

[1] Thymine
[2] Guanine
[3] Nucleic Acid

ـ تکه‌تکه‌شدن DNA: این رویکرد برای ساخت کتابخانه‌ای از توالی DNA به‌کار می‌رود. همچنین، برای تقسیم توالی DNA به قطعات کوچک استفاده می‌شود. بسیاری از الگوریتم‌های رمزگذاری، از این روش به‌عنوان لایه دوم امنیتی بهره می‌برند. این تکنیک در رمزگذاری کلید نیز پیاده‌سازی شده است [۲۶,۲۹].

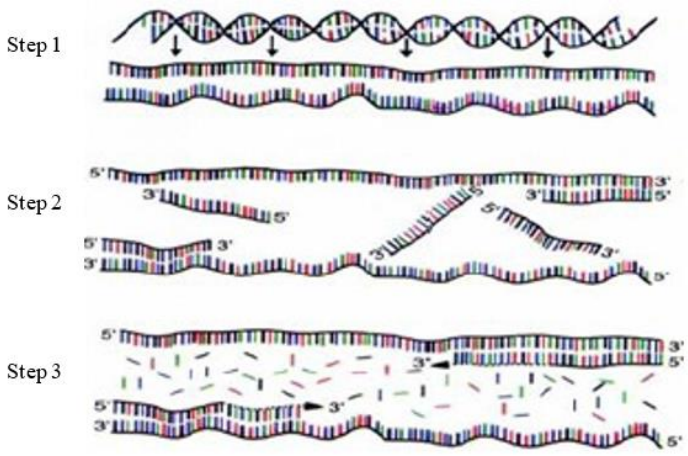
ـ کدگذاری دیجیتالی DNA: این تکنیک، نقش حیاتی در رمزگذاری و رمزگشایی اطلاعات ایفا می‌کند که از آن برای نگاشت بازهای DNA به‌صورت ارقام دودویی 0 و 1 استفاده می‌شود. این فناوری، بازهای A، G، C، T را به‌ترتیب با 00، 10، 01، 11 نشان می‌دهد. همچنین، راه‌کاری برای انتقال مقادیر دودویی به‌وسیله بازها را فراهم می‌کند [۲۶]. پیام‌های متنی ساده را می‌توان به‌راحتی با استفاده از این طرح کدگذاری کرد. به‌طور مثال، شخصی مایل است از کدگذاری DNA برای انتقال عدد ۹۷ استفاده کند. برای شروع این فرآیند، عدد ۹۷ به‌صورت دودویی تبدیل می‌شود. با تبدیل ارقام دودویی ۹ و ۷ به شکل‌های چهار بیتی، به ترتیب 1001 و 0111 حاصل می‌شود. سپس، نمایش‌های ارقام دودویی ۹ و ۷ باهم ترکیب می‌شود. نتیجه، عدد دودویی 10010111 خواهد بود. سپس، این رشته دودویی به‌صورت توالی DNA کدگذاری می‌شود. درنهایت، پیام کدگذاری‌شده به‌صورت "GCCT" به گیرنده مورد نظر تحویل داده می‌شود. این روش کدگذاری، مبنای الگوریتم‌های رمزگذاری برای رویکردهای کدگذاری دیجیتالی DNA است. همچنین به‌دلیل محاسبات زیستی و ماهیت امنیتی الگوریتم، کاربرد وسیع‌تری دارد [۲۹].

ـ واکنش زنجیره‌ای پلیمراز: این یک فرآیند برای ایجاد کپی‌های زیاد از توالی‌های DNA است. هدف از این کار، کدگذاری پیام بین دو پرایمر است. پرایمر، یک توالی اسید نوکلئیک با طول کوتاه است که نقطه شروع سنتز DNA را فراهم می‌کند. آنزیم‌هایی که DNA را سنتز می‌کنند، DNA پلیمراز نامیده می‌شوند. DNA پلیمرازها، تنها قادر به متصل‌کردن رشته نوکلئوتیدهای موجود با نوکلئوتیدهای جدید هستند. بنابراین، پرایمر باید تحت فرآیند سنتز قرار گیرد؛ به‌گونه‌ای که پرایمر به‌عنوان کلید و پایه سنتز DNA به‌کار رود. درنتیجه، وقتی توالی کدگذاری‌شده تقویت می‌شود که جفت‌های پرایمر PCR ناشناخته باشند [۶,۱۰]. مطابق با شکل ۲، مراحل عمل بیولوژیکی PCR به شرح زیر است [۲۶]:

مرحله اول: PCR با فرآیند ساختارشکنی[1] آغاز می‌شود که در آن، یک DNA دو رشته‌ای از هم جدا شده و به دو تک‌رشته DNA تبدیل می‌شود. این کار با حرارت‌دادن DNA الگو در دمای بسیار بالا در حدود ۹۴ تا ۹۶ درجه سانتی‌گراد انجام می‌شود.

مرحله دوم: این مرحله برای پردازش اتصال پرایمر به تک‌رشته DNA است. در این‌جا، دما بین ۵۰ تا ۶۵ درجه سانتی‌گراد خنک می‌شود تا پرایمرهای مربوطه به دنباله‌های مکمل خودشان متصل شوند. ماهیت پرایمرها، تکثیر DNA مجاور است.

مرحله سوم: این مرحله برای بسط پرایمر است. در این‌جا، درجه حرارت به ۷۲ درجه سانتی‌گراد افزایش می‌یابد تا آنزیم پلیمراز شروع به فعالیت کند. آنزیم پلیمراز، بازهای نوکلئوتیدی را به ادامه پرایمرها به‌روی رشته DNA اصلی اضافه می‌کند. رشته‌های DNA بین پرایمرها تکثیر می‌شوند.



شکل ۲: واکنش زنجیره‌ای پلیمراز [۳۱]

Figure 2. Polymerase chain reaction [31]

[1] Denaturation

ـ هیبریدسازی DNA: توالی DNA دارای دو جفت باز (A -T) و (C - G) به‌همراه یک گروه قند و فسفات است. ترکیب این دو جفت‌باز، ساختار مارپیچ‌گونه DNA را تشکیل می‌دهد. توالی این بازها، اطلاعات موجودات زنده را ذخیره می‌کند؛ به‌طوری‌که برای همه موجودات زنده منحصر به فرد است. بازهای A و T مکمل هم هستند و بازهای C و G نیز باهم مکمل‌اند؛ جفت‌شدن این بازها با مکمل خودشان، همان فرآیند هیبریدسازی است. شکل ۳ این فرآیند را نشان می‌دهد [۶].

شکل ۳: فرآیند هیبریدسازی دو تک‌رشته DNA باهم [۲۷]

Figure 3. The process of hybridizing two single-stranded DNAs together [27]

جدول ۱: انواع مختلف فناوری‌های DNA

Table 1. Different types of DNA technologies

DNA technology	Main specifications
One-Time Pad (OTP)	• A key is randomly generated using OTP. • The key generated by OTP can only be used once for encryption and changes for the next use. • The key size is greater than or equal to the plaintext message. • The key must be defined as unique and remain secret. • Guessing the correct key to obtain the plaintext in OTP technique is difficult.
DNA Chip Technology	• DNA chip technology is useful for storing and managing biological information. • DNA chips are also known as gene chips and microarrays. • DNA chip characteristics may change due to some environmental factors and negatively affect the chiphertext.
DNA Fragmentation	• DNA fragmentation is necessary for building a library for DNA sequencing. • DNA fragmentation is used to divide DNA sequences into smaller fragments. • DNA fragmentation can be utilized as a second security layer or for key encryption.
Polymerase Chain Reaction (PCR)	• PCR is a DNA amplification and quantification technique. • Two primers (small DNA fragments) can be used as the main key for the PCR amplification process. • To encode original data in a new sequence, the original information can easily be placed between two primers. • Primer accuracy in this process is very important because primers of different lengths provide different results.
DNA Digital Encoding	• DNA digital encoding is used to convert binary strings to DNA sequences and vice versa. • DNA digital encoding is the most common technology used in DNA-based cryptographic systems.
DNA Hybridization	• In the hybridization method, two single-strand DNAs are combined to produce a single DNA sequence. • Using DNA hybridization is recommended to minimize time complexity in DNA cryptographic systems.

۴- روش پیشنهادی

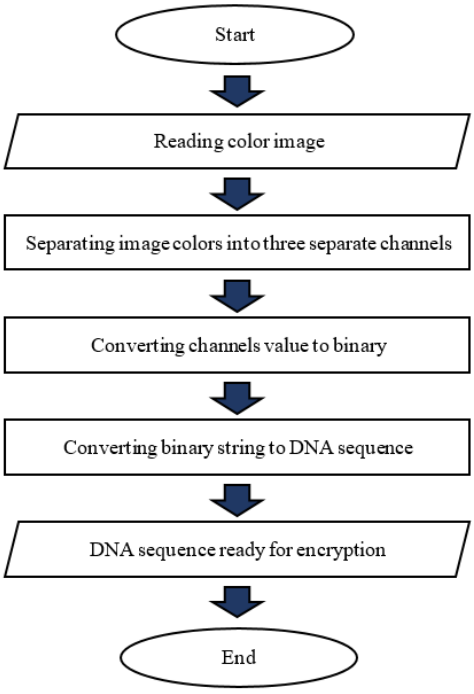
اگرچه روش‌هایی برای رمزنگاری DNA ارائه شده است؛ ولی، بیشتر این روش‌ها روی متن‌های انگلیسی تمرکز دارند و کاری بر روی متن‌های فارسی انجام نشده است. مفهوم استفاده از DNA برای رمزنگاری متن‌های فارسی، هنوز مراحل اولیه را می‌پیماید. بنابراین به دنبال الگوریتمی جدید هستیم که بتوان با استفاده از آن، متن‌های فارسی را رمزنگاری کرد؛ به‌طوری‌که این الگوریتم بر مبنای DNA باشد. از سوی دیگر، تصاویر به‌عنوان رسانه‌ای رایج به‌خصوص در اینترنت شناخته می‌شوند. از این‌رو، راه‌کاری برای رمزگذاری تصاویر با استفاده از این طرح نیز پیشنهاد شده است.

اگر داده‌ها به‌صورت متن باشند، با استفاده از این الگوریتم به‌راحتی رمزگذاری می‌شوند. ولی، چنان‌چه داده‌ها به‌صورت تصویر باشند، ابتدا تصویر به توالی DNA تبدیل می‌شود. سپس، از همین روش رمزگذاری متن استفاده می‌شود. روش پیشنهادی از فازهای تولید کلید راز، رمزگذاری و رمزگشایی تشکیل‌شده‌است. ولی، اگر تصمیم بر رمزنگاری تصویر باشد، از یک فاز اضافی با عنوان پیش‌پردازش تصویر استفاده می‌شود. در ادامه به بررسی هر یک از این فازها پرداخته می‌شود.

۴-۱- فاز پیش‌پردازش برای تصویر

همان‌طور که پیشتر بیان شد، این فاز مختص داده‌های تصویری است. در این فاز، از تابع شدت رنگ برای هر پیکسل استفاده می‌شود تا تصویر به‌صورت رشته دودویی درآید. بدین‌ترتیب می‌توان تصویر را به سه کانال رنگی قرمز، سبز و آبی (RGB)[1] تقسیم کرد؛ چراکه RGB رنگ‌های اصلی هستند که سایر رنگ‌ها نیز از روی این سه رنگ ساخته می‌شوند. شدت هر کانال رنگی را به‌طور معمول می‌توان با استفاده از هشت بیت ذخیره کرد. یعنی هر پیکسل در تصویر رنگی به فضای ۲۴ بیتی نیاز دارد؛ سه کانال رنگی وجود دارد که هر کدام، هشت بیت فضا اشغال می‌کند (۲۴ = ۸ × ۳). اکنون به‌راحتی می‌توان، مقدار هر کانال رنگی را به‌صورت دودویی تبدیل کرد. درنهایت، یک رشته دودویی به‌دست می‌آید که می‌توان آن‌را با استفاده از فناوری کدگذاری دیجیتالی DNA به یک توالی DNA تبدیل کرد. سایر مراحل به‌منظور رمزنگاری‌کردن این توالی DNA مطابق با سایر فازها انجام می‌شود.

جالب این‌جا است که داده‌های دودویی به‌دست‌آمده از تصویر، در قالب یک فایل متنی رمزگذاری می‌شود. بنابراین، طرح پیشنهادی دارای ویژگی اضافه‌ای برای پنهان‌سازی تصاویر به‌صورت فایل متنی است. شکل ۴ فلوچارت مراحل فاز پیش‌پردازش تصویر را نشان می‌دهد.



شکل ۴: فلوچارت مراحل فاز پیش‌پردازش تصویر
Figure 4. Flowchart of image pre-processing phase stages

۴-۲- فاز تولید کلید راز

در این فاز، از سیستم HECC به‌منظور تولید کلیدهای عمومی و خصوصی موجودیت‌ها استفاده می‌شود تا بدین‌وسیله بتوان کلید راز انتخابی توسط فرستنده را رمز کرده و روی کانال ناامن فرستاد. بنابراین، دیگر نیازی به کانال امن به‌منظور اشتراک‌گذاری کلید راز نیست. رمزنگاری منحنی فرابیضوی [۳]، یکی از سیستم‌های کلید عمومی است که در گره‌های سیار با منابع محدود به‌منظور دستیابی به بهینه‌سازی انرژی و امنیت بیشتر استفاده می‌شود. همچنین، سربار محاسباتی پایینی دارد و می‌تواند به‌راحتی برای کاربردهای سخت‌افزاری و نرم‌افزاری پیاده‌سازی شود. این روش، کارایی (سرعت) بالایی دارد. مراحل تولید کلید به‌صورت زیر است:

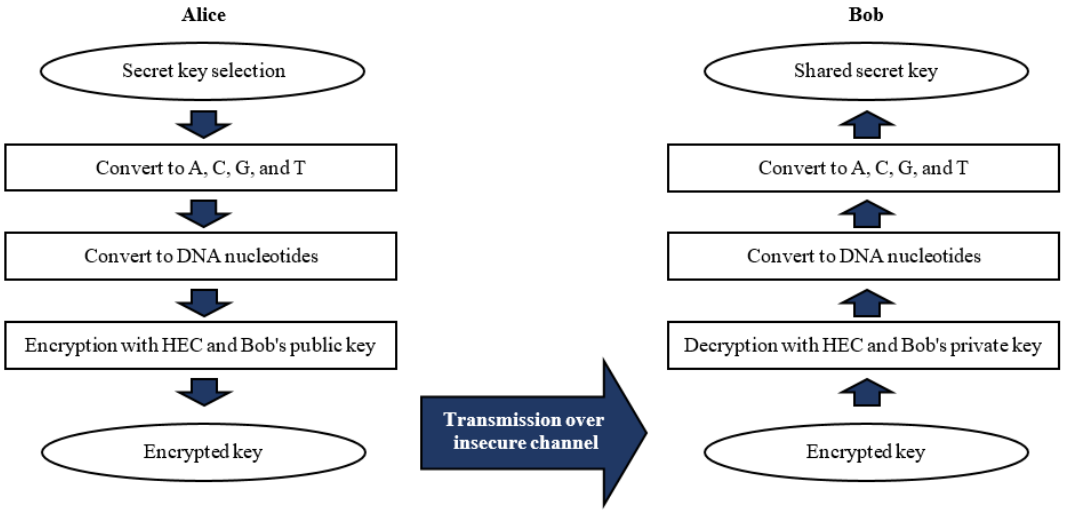
[1] Red Green Blue

مرحله ۱: در این مرحله، پارامترهای اصلی موجودیت‌های آلیس (فرستنده) و باب (گیرنده) برای انجام عملیات رمزنگاری تولید می‌شوند. برای این منظور، منحنی C مورد استفاده قرار می‌گیرد. پارامترهای منحنی C، عدد اول P، مقسوم‌علیه D و پارامترهای عمومی هستند که در دسترس همگان قرار می‌گیرند. جدول ۲ علائم مورد استفاده در الگوریتم پیشنهادی را بیان می‌کند.

مرحله ۲: آلیس، یک خم HEC را از جنس g ($g > 2$) روی میدان F انتخاب می‌کند تا رابطه $C: y^2 + h(x)\, y = f(x)$ حاصل شود.

جدول ۲: علائم به‌کار رفته در الگوریتم پیشنهادی
Table 2. Symbols used in the proposed algorithm

Symbol	Description
$\|\|$	Concatenation operator
XOR	An XOR operation
A_S	Alice's private key
B_S	Bob's private key
P_A	Alice's public key
P_B	Bob's public key
C	Curve of type HEC
F	Finite field
G	Point of order n ($n \times G = 0$)



شکل ۵: فرآیند رمزکردن کلید راز و به اشتراک‌گذاری آن
Figure 5. Process of encrypting and sharing the secret key

مرحله ۳: آلیس، مجموعه‌ای از عناصر ژاکوبین[1] بر روی $J(F_q)$ را با استفاده از نقاط خم منحنی C ایجاد می‌کند که می‌تواند به‌صورت $D = \sum m_i P_i$ نوشته شود. در این‌جا، D مقسوم‌علیه است، $m \geq 0$ است و P_i نقطه متناهی است.

مرحله ۴: آلیس، نقطه G از درجه n را تولید می‌کند؛ به‌طوری‌که $n \times G = 0$ شود. سپس، عدد تصادفی A_S را به‌عنوان کلید خصوصی خود انتخاب می‌کند. همچنین، کلید عمومی P_A را به‌صورت $P_A = A_S \times G$ محاسبه می‌نماید.

مرحله ۵: همانند مرحله ۳، باب نیز یک عدد تصادفی B_S را به‌عنوان کلید خصوصی خود انتخاب کرده و $P_B = B_S \times G$ را محاسبه می‌کند. بنابراین، P_B کلید عمومی موجودیت باب خواهد بود.

مرحله ۶: آلیس در نهایت، پارامترهای (P_A, G, D, C, n) را منتشر می‌کند.

[1] Jacobian

هر دو موجودیت آلیس و باب دارای یک‌جفت کلید (خصوصی و عمومی) هستند. بنابراین، می‌توانند کلید راز را به‌صورت یک متن ساده در نظر بگیرند و با استفاده از تکنیک HECC آن‌را رمز کرده و با یک‌دیگر به اشتراک بگذارند. برای این منظور، آلیس یک کلید راز را انتخاب می‌کند. سپس، آن‌را با کلید عمومی باب رمز کرده و از طریق یک کانال ناامن برای باب می‌فرستند. باب پس از دریافت کلید رمزشده، با استفاده از کلید خصوصی خود، کلید را رمزگشایی می‌کند و به کلید راز انتخابی از طرف آلیس دست می‌یابد. از این پس، مراحل رمزگذاری و رمزگشایی پیام‌ها با استفاده از این کلید راز اشتراکی انجام می‌شود. فرآیند رمزکردن و اشتراک‌گذاری کلید راز در شکل ۵ نشان داده‌شده‌است.

۴-۳- فاز رمزگذاری

این فاز شامل مراحل زیر است:

مرحله ۱: ابتدا، متن ساده فارسی به DNA کدشده تبدیل می‌شود. برای انجام این کدگذاری، از جدول ۳ استفاده می‌شود. کدهای DNA فقط برای حروف فارسی، ارقام و علائم نگارشی استفاده می‌شوند.

جدول ۳: حروف و بازهای DNA معادل با آن‌ها
Table 3. DNA bases and equivalent letters

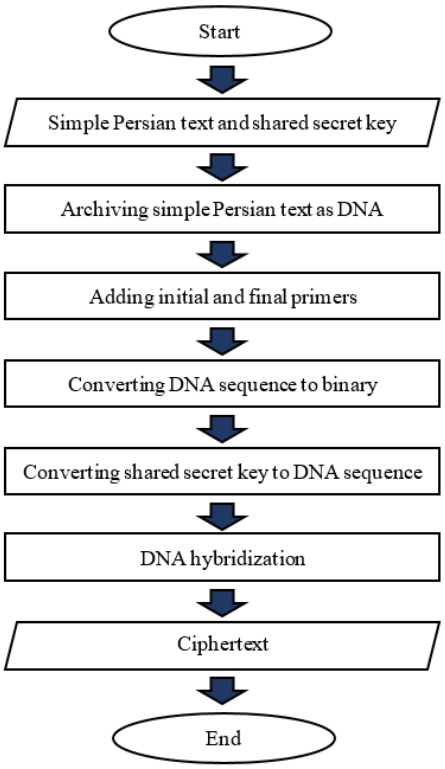
DNA Base	Letter	DNA Base	Letter	DNA Base	Letter
GAAC	؛	CGAG	ش	CTCG	آ
ATAT	:	ATCT	ص	CGTC	أ
CACC	)	CCCA	ض	GATC	إ
ACTC	(	CCAC	ط	CGTT	ء
CGGG	»	ATTA	ظ	TGCG	ئ
TTCT	«	CTCC	ع	GTAA	ؤ
TCAC	...	CCGA	غ	GCAT	ا
CAAT	؟	AAAG	ف	AATC	ب
TATG	!	CAAA	ق	ATTG	پ
ATAG	_	AATT	ک	GGCA	ت
TGTG	Space	ATGA	گ	CCGC	ث
GCCT	٠	GGGC	ل	GACG	ج
CTAT	۱	TAAG	م	AGAA	چ
TCAT	۲	TCGG	ن	AGTA	ح
ACGG	۳	TGCT	و	GCGA	خ
CTTG	۴	TTAG	ه	GAGG	د
AGGT	۵	GTGA	ی	TGGC	ذ
TGAG	۶	TACA	ّ	AATA	ر
TGCC	۷	CATG	ً	CCTA	ز
CGCC	۸	TCTT	.	CCAT	ژ
CTTA	۹	TCCG	،	TCTC	س

مرحله ۲: بیت‌های اضافی (یعنی، پرایمرهای ابتدایی و انتهایی) به خروجی مرحله قبل اضافه می‌شود.

مرحله ۳: کلید راز اشتراکی که به‌صورت DNA کدشده است، به اندازه طول پیام تکرار می‌شود.

مرحله ۴: رشته DNA حاصل از مرحله ۳ به رشته هم‌ارز دودویی تبدیل می‌شود. برای این تبدیل، از فناوری کدگذاری دیجیتالی DNA استفاده می‌شود.

مرحله ۵: به‌منظور هیبریدسازی DNA، نیاز است که طول رشته هم‌ارز دودویی به اندازه طول کلید (مثلاً ۶۴۰ بیت) شود که این ۶۴۰ بیت با استفاده از چسباندن توالی تکراری به‌دست می‌آید. در این مرحله، هیبریدسازی DNA انجام می‌شود. به‌عنوان ورودی، توالی DNA حاصل از کلید راز اشتراکی و خروجی کدشده هم‌ارز دودویی را دریافت می‌کند. به‌منظور هیبریدسازی DNA، اگر بیت دودویی حاصل از رشته هم‌ارز دودویی "صفر" باشد، عملی انجام نمی‌شود؛ در صورتی که مقدار این بیت برابر با "یک" باشد، آن‌گاه بازهای 10 متناظر آن در رشته دودویی حاصل از کلید راز اشتراکی مکمل می‌شود و سپس، به‌صورت DNA نوشته می‌شود. شکل ۶ فلوچارت مراحل فاز رمزگذاری را نشان می‌دهد.



شکل ۶: فلوچارت مراحل فاز رمزگذاری
Figure 6. Flowchart of encryption phase stages

۴-۳-۱- مثال کاربردی

یک مثال کاربردی برای توضیح الگوریتم پیشنهادی به‌صورت زیر بیان می‌شود:

متن ساده "موش" را در نظر بگیرید. ابتدا، یک توالی DNA تصادفی از پایگاه داده شورای ملی اطلاعات بیوتکنولوژی (NCBI)[1] از اورگانیسم DOG انتخاب می‌شود. این پایگاه داده به‌صورت عمومی و در دسترس همگان است. نام علمی این اورگان، گونه کانیس لوپوس[2] است. این توالی نوکلئوتیدها، برای تعیین پرایمرهای ابتدایی و انتهایی استفاده می‌شوند. کلمه "موش" با استفاده از جدول ۲ به‌صورت DNA کد می‌شود. سپس، پرایمرهای ابتدایی و انتهایی به ابتدا و انتهای آن اضافه می‌شود. بدین ترتیب، توالی DNA برای متن ساده "موش" به‌صورت شکل ۷ درمی‌آید.

موش	→	م	و	ش
		TAAG	TGCT	CGAG

شکل ۷: توالی DNA مربوط به متن ساده موش
Figure 7. DNA sequence related to plain text of mouse

[1] National Council of Biotechnological Information
[2] Canis Lupus Familiaris

پس از اضافه‌شدن پرایمرهای ابتدایی و انتهایی به توالی بالا، شکل ۸ حاصل می‌شود.

GTCCACATCA	TAAG TGCT CGAG	GCAGCAGCCC
20 bit	24 bit	20 bit
64 bit		

شکل ۸: توالی DNA همراه با پرایمرها
Figure 8. DNA sequence with primers

در این مثال برای راحتی کار، پرایمرهای ابتدایی و انتهایی به اندازه ۲۰ بیت و با طول ثابت درنظر گرفته‌شده‌است. اکنون، توالی DNA همراه با پرایمرها براساس فناوری کدگذاری دیجیتالی DNA به رشته دودویی معادل تبدیل‌شده و به‌صورت شکل ۹ نمایش داده می‌شود.

G T C C A C A T C A	T A A G	T G C T	C G A G	G C A G C A G C C C
10 11 01 01 00 01 00 11 01 00	11 00 00 10	11 10 01 11	01 10 00 10	10 01 00 10 01 00 10 01 01 01

شکل ۹: رشته دودویی از توالی DNA نهایی
Figure 9. Binary string from final DNA sequence

اکنون، کلید راز اشتراکی به‌شکل توالی DNA هم‌ارز تبدیل می‌شود. به‌عنوان مثال، کلید راز اشتراکی به‌صورت شکل ۱۰ در نظر گرفته‌شده‌است. ولی، برای هیبریدسازی DNA لازم است که طول توالی DNA به‌دست‌آمده از کلید راز اشتراکی به طول ۶۴۰ بیت برسد. با چسباندن تکراری این توالی، طول ۶۴۰ بیتی تولید می‌شود. در نهایت، کلید راز اشتراکی همانند شکل ۱۱ می‌شود.

GCATTAACAGAGAGGA (32 bit)

شکل ۱۰: نمونه کلید راز اشتراکی
Figure 10. Sample of shared secret key

GCATTAACAGAGAGGAGCATTAACAGAGAGGAGCATTAACAGAGAGGAG
CATTAACAGAGAGGAGCATTAACAGAGAGGAGCATTAACAGAGAGGAGCA
TTAACAGAGAGGAGCATTAACAGAGAGGAGCATTAACAGAGAGGAGCAT
TAACAGAGAGGAGCATTAACAGAGAGGAGCATTAACAGAGAGGAGCA
TTAACAGAGAGGAGCATTAACAGAGAGGAGCATTAACAGAGAGGAGC
ATTAACAGAGAGGAGCATTAACAGAGAGGAGCATTAACAGAGAGGAG
CATTAACAGAGAGGAGCATTAACAGAGAGGA

شکل ۱۱: کلید راز اشتراکی برای هیبریدسازی
Figure 11. Shared secret key for hybridization

به‌منظور هیبریدسازی DNA، اگر بیت دودویی حاصل از رشته ورودی و پرایمرها صفر باشد، عملی انجام نمی‌شود؛ ولی اگر یک باشد، آن‌گاه بازهای 10 متناظر آن در رشته دودویی به‌دست‌آمده از کلید راز اشتراکی، مکمل خواهد شد و سپس، به‌صورت DNA نوشته می‌شود. رشته DNA حاصل از هیبریدسازی، به‌صورت متوالی تکرار می‌شود تا طول ۶۴۰ بیت تولید شود. بنابراین، متن رمز شده به‌صورت شکل ۱۲ به‌دست می‌آید که در نهایت، شکل ۱۳را خواهیم داشت.

Input binary string and primers	→	**1**	0	**1**	**1**	0	**1**	0	**1**	⋯
Shared secret key	→	G	C	A	T	T	A	A	C	⋯
		↓	↓	↓	↓	↓	↓	↓	↓	
Binary conversion	→	10	01	00	11	11	00	00	01	⋯
		↓		↓	↓		↓		↓	
Complement of ones bits	→	**01**	01	**11**	**00**	11	**11**	00	**10**	⋯
		↓	↓	↓	↓	↓	↓	↓	↓	
Ciphertext (640-bits)	→	C	C	T	A	T	T	A	G	⋯

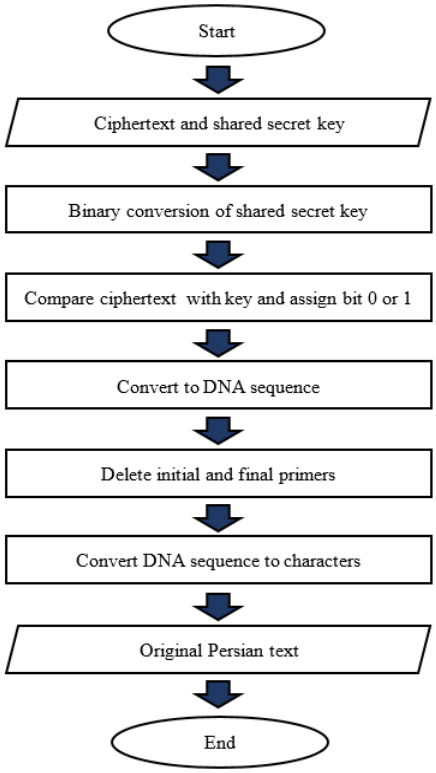
شکل ۱۲: نحوه به‌دست‌آوردن متن رمزی
Figure 12. How to obtain ciphertext

CCTATTAGAGACAGCTGGATATACAGTGTCCAGGTATTTCAGTGTGGTGC
TTTTACTGACACGTCCTATTAGAGACAGCTGGATATACAGTGTCCAGGTATTT
CAGTGTGGTGCTTTTACTGACACGTCCTATTAGAGACAGCTGGATATACAGT
GTCCAGGTATTTCAGTGTGGTGCTTTTACTGACACGTCCTATTAGAGACAGCT
GGATATACAGTGTCCAGGTATTTCAGTGTGGTGCTTTTACTGACACGTCCTAT
TAGAGACAGCTGGATATACAGTGTCCAGGTATTTCAGTG

شکل ۱۳: متن رمزی نهایی
Figure 13. Final ciphertext

۴-۴- فاز رمزگشایی

شکل ۱۴ فلوچارت مراحل فاز رمزگشایی را نشان می‌دهد.



شکل ۱۴: فلوچارت مراحل فاز رمزگشایی
Figure 14. Flowchart of decryption phase stages

به‌منظور رمزگشایی، متن رمزی به‌همراه کلید راز اشتراکی دریافت می‌شوند؛ به‌طوری‌که هر دو، ۶۴۰ بیتی و به‌صورت توالی DNA هستند. سپس، مراحل زیر دنبال می‌شوند:

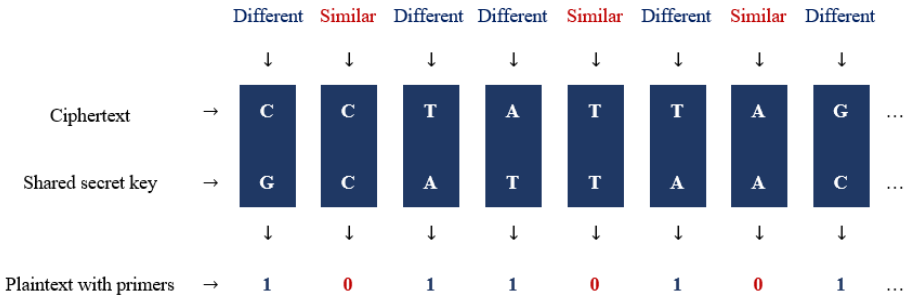
مرحله ۱: متن رمزی و کلید راز اشتراکی به‌صورت رشته‌های دودویی نوشته می‌شوند. هر جایی که دو بیت از متن رمزشده و دوبیت از کلید راز اشتراکی یکسان بودند، بیت صفر متناظر در خروجی تولید می‌شود و چنان‌چه یکسان نبودند، بیت یک در خروجی نوشته می‌شود.

مرحله ۲: پس از این‌که رشته متن اصلی به‌صورت دودویی به‌دست آمد، به‌شکل یک توالی DNA تبدیل می‌شود.

مرحله ۳: گیرنده، پرایمرهای ابتدایی و انتهایی را حذف می‌کند و آن‌چه باقی می‌ماند را مطابق با جدول ۲ به حروف فارسی متناظر تبدیل می‌کند.

۴-۴-۱- مثال کاربردی

متن رمز شده در شکل ۱۳را در نظر بگیرید. با توجه به فازهای ذکر شده برای رمزگشایی، همانند شکل ۱۵ عمل می‌کنیم تا به رشته دودویی متن اصلی به‌همراه پرایمرها دست پیدا کنیم.



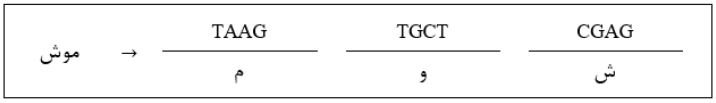
شکل ۱۵: مراحل رسیدن به رشته دودویی متن اصلی همراه با پرایمرها
Figure 15. Steps to get binary string of plaintext with primers

اکنون، رشته دودویی حاصل در شکل ۱۵ با استفاده از فناوری کدگذاری دیجیتالی DNA به‌صورت توالی DNA تبدیل می‌شود. این توالی در شکل ۱۶ نشان داده‌شده‌است.

10 11 01 01 00 01 00 11 01 00	11 00 00 10	11 10 01 11	01 10 00 10	10 01 00 10 01 00 10 01 01 01
G T C C A C A T C A	T A A G	T G C T	C G A G	G C A G C A G C C C

شکل ۱۶: توالی DNA مربوط به متن اصلی و پرایمرها
Figure 16. DNA sequence related to plaintext and primers

شکل ۱۶ توالی DNA مربوط به متن اصلی و پرایمرها است که گیرنده، پرایمرهای ابتدایی و انتهایی را از آن حذف می‌کند تا تنها، توالی DNA مربوط به متن اصلی باقی بماند. سپس، با استفاده از جدول ۲ می‌تواند به متن ساده "موش" برسد. این فرآیند در شکل ۱۷ قابل مشاهده است.



شکل ۱۷: رسیدن به متن ساده موش از روی توالی DNA
Figure 17. Getting plaintext of mouse from DNA sequence

۵- ارزیابی روش پیشنهادی

در این بخش، سیستم رمز پیشنهادی از نظر امنیت و کارایی مورد بررسی و ارزیابی قرار می‌گیرد. محیط ارزیابی طرح پیشنهادی معرفی می‌شود. سپس، به اجرای آن پرداخته می‌شود و در آخر، عملکرد روش پیشنهادی با روش‌های پیشین مقایسه می‌شود.

۵-۱- تحلیل امنیتی

در عصر حاضر، احتمال وقوع هرگونه خطر امنیتی باید به‌طور مؤثری در نظر گرفته شود. همچنین، فراموش نکنیم که یک مهاجم DNA، اصول فناوری DNA را می‌داند. بنابراین، هرچه از فناوری پیشرفته‌تری استفاده شود، هم‌چنان خطر امنیتی بزرگی را به‌همراه خواهد داشت. یک راه‌حل برای تأمین امنیت سیستم را می‌توان در تولید کلید و روند به اشتراک‌گذاری آن پیدا کرد. ویژگی‌های امنیتی طرح پیشنهادی به شرح زیر است:

۱- در سیستم پیشنهادی، کلید راز اشتراکی برای رمزگذاری و رمزگشایی مورد استفاده قرار می‌گیرد که تبادل آن برخلاف دیگر روش‌ها به یک کانال امن نیازی ندارد؛ چراکه با استفاده از تکنیک HECC به اشتراک گذاشته می‌شود که در بخش قبلی بیان شد.

۲- اندازه کلید، یکی از مهم‌ترین عوامل برای امن‌کردن سیستم رمزنگاری است. یک سیستم رمزنگاری کارآمد باید کلیدی داشته باشد که به اندازه کافی بزرگ باشد تا امنیت آن تضمین شود. اگر فضای کلید به اندازه کافی بزرگ نباشد، سیستم رمز در برابر حمله جستجوی فراگیر آسیب‌پذیر می‌شود. از این‌رو طرح پیشنهادی نسبت به سایر طرح‌های مشابه، فضای کلید بزرگتر (۶۴۰ بیتی) را ارائه داده است.

امنیت سیستم پیشنهادی دارای دو سطح است:

۱- سطح اول، امنیت اطلاعات بیولوژیکی و پیچیدگی مسائل سخت (هیبریدسازی DNA) است. بیت‌های اضافی (پرایمرهای ابتدایی و انتهایی) در توالی اصلی DNA، یک توالی جعلی DNA را ایجاد می‌کند. این توالی با کلید راز اشتراکی ترکیب می‌شود تا متن رمزشده به‌دست آید. این عمل، منجر به آسیب‌پذیری کم‌تر متن ساده در برابر مهاجمان می‌شود. در واقع، پیش‌بینی متن اصلی رمزشده برای مهاجمان دشوارتر می‌شود؛ زیرا متن رمزشده، حاوی اطلاعات اضافی در قالب پرایمر ابتدایی و انتهایی است. مهاجمان نیاز دارند که برای تشخیص متن رمز شده اصلی، هر دو پرایمر را بدانند. البته، اگر آن‌ها به پرایمرها دست پیدا کنند، امنیت سیستم رمزنگاری هنوز دست‌نخورده باقی می‌ماند.

۲- سطح دوم امنیت، مسئله محاسبات پیمانه‌ای در تکنیک HEC است. بدون دانستن کلید خصوصی موجودیت‌ها، اخلال‌گر نمی‌تواند کلید راز اشتراکی را محاسبه کند. بنابراین، بدون اطلاع از کلید راز اشتراکی، اخلال‌گر باید محاسبات دشواری را برای شکستن امنیت سیستم رمز پیشنهادی انجام دهد.

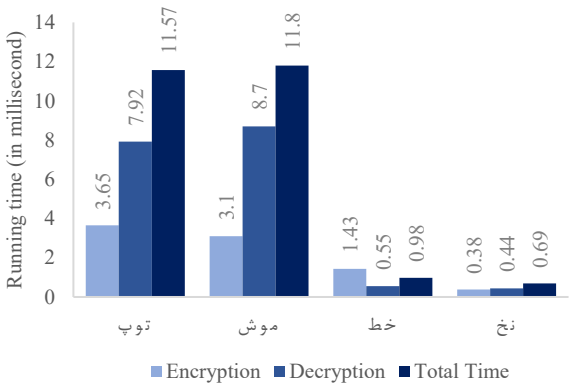
با توجه به آن‌چه گفته شد، سیستم رمز پیشنهادی به‌دلیل سختی مسائل بیولوژیکی و سختی محاسبه کلید راز اشتراکی، امنیت دو چندانی را فراهم کرده است.

۵-۲- ارزیابی کارایی

در این بخش، پروتکل پیشنهادی بر اساس نتایج پیاده‌سازی با استفاده از زبان C# مورد ارزیابی قرار می‌گیرد. مشخصات سیستم به‌کار رفته برای پیاده‌سازی در جدول ۴ آمده است. پنجره خروجی برنامه به ازای متن ساده فارسی "موش" در شکل ۱۸ مشخص شده است. همان‌طور که شکل ۱۸ نشان می‌دهد، درستی عملکرد الگوریتم پیشنهادی اثبات می‌شود. همچنین، در شکل ۱۹ زمان اجرای الگوریتم پیشنهادی به ازای چند متن ساده فارسی مشخص شده است. براساس نتایج پیاده‌سازی، هر چه تعداد کاراکترها کم‌تر باشد، این طرح با سرعت بیشتری اجرا می‌شود.

جدول ۴: مشخصات سیستم به‌کار رفته جهت پیاده‌سازی
Table 4. Specifications of system used for implementation

Hardware/Software	Specifications
Operating System	Windows 11
Main Memory (RAM)	32GB
Processor (CPU)	Intel Core i7-10750 M (2.60 GHZ)



شکل ۱۹: مقایسه زمان اجرای طرح پیشنهادی برای چند نمونه از متون فارسی
Figure 19. Comparison of execution time of proposed scheme for several Persian text samples

شکل ۱۸: پنجره خروجی برنامه به ازای متن ساده موش
Figure 18. Program output window for plaintext of mouse

دیدیم که در این روش، رمزگذاری متون فارسی با استفاده از کلید راز اشتراکی انجام می‌شود؛ به‌طوری‌که برای تبادل کلید راز اشتراکی از سیستم رمز منحنی فرابیضوی استفاده می‌شود. روش‌های تبادل کلید، نیاز به کانال امن برای به اشتراک‌گذاری کلید راز را برطرف می‌کنند. جالب این‌جاست که در مرجع [۳] اثبات شده است که HECC نسبت به سایر روش‌های تبادل کلید، سربار محاسباتی کم‌تر و سرعت بیشتری دارد. همان‌طور که جدول ۵ نشان می‌دهد، سیستم رمز مبتنی‌بر DNA پیشنهادی با ترکیب‌کردن روش‌های رمزنگاری متقارن و نامتقارن، از مزایای هر دو روش بهره‌مند شده و عملکرد بهتری دارد.

۶- نتیجه‌گیری

نگرانی‌ها در مورد محرمانگی اطلاعات، منجر به استفاده از الگوریتم‌های رمزنگاری شده است. به‌همین منظور، الگوریتم‌های امنیتی به دو دسته عمده متقارن و نامتقارن تقسیم شدند و مورد بررسی قرار گرفتند. الگوریتم‌های متقارن، بسیار سریع و امن هستند. با این حال، با روی کار آمدن کامپیوترهای کوانتومی، این روش‌های رمز شکسته می‌شوند. از این‌رو، روش‌های رمزنگاری DNA امید تازه‌ای را در برابر حمله‌های کوانتومی به‌وجود آورده‌اند. فناوری DNA دارای مزایای منحصر به‌فردی نسبت به رمزنگاری سنتی است؛ به‌ویژه این‌که مصرف انرژی پایین و ظرفیت ذخیره‌سازی بالایی دارد. در میان سیستم‌های رمز DNA،

تکنیک‌هایی که مبتنی بر هیبریدسازی هستند، سرعت اجرای بالاتری دارند. از سوی دیگر، به‌کار بردن روش‌هایی برای به‌دست آوردن کلید راز اشتراکی بدون نیاز به انتقال کلید روی کانال امن، باعث کاهش سربار محاسباتی شده و کارایی بیشتری را فراهم می‌کنند. طبق بررسی‌های انجام‌شده، تاکنون هیچ یک از روش‌های رمزنگاری، از تمامی این ویژگی‌ها برای رمزگذاری متون فارسی و تصاویر بهره‌ای نبرده است.

جدول ۵: ارزیابی کارایی الگوریتم‌های مبتنی‌بر DNA
Table 5. Performance evaluation of DNA-based algorithms

Algorithm	Sample Text	Method	Performance Evaluation	Time Complexity
An Efficient Hybrid Elliptic Curve Cryptosystem with DNA Encoding [12]	Letter t	Asymmetric	Lower efficiency compared to HECC - exchange method, additional information, delayed increase, and high cost (due to performing encryption with the public key)	$O(n^2)$
DNA Cryptography Based on Symmetric Key Exchange [6]	1. SE 2. Hello 3. DNA	Symmetric	Need for secure channel (due to using symmetric key exchange)	O(n)
An Enhanced Polyalphabetic Algorithm on Vigenerecipher with DNA Cryptography [15]	Letter A	Symmetric	High computational overhead (due to using calculations and predictive operations)	O(n)
A Novel Level-Based DNA Security Algorithm Using DNA Codons [16]	Desire	Symmetric	Low execution speed (due to performing multiple repetitive operations)	O(n)
Designing Cloud Security Using Multi-layer DNA Cryptography in Python [19]	Letter A	Symmetric	Increased communication overhead (due to processing data for cloud)	O(n)
A New DNA Cryptosystem for Encrypting Persian Texts and Images with Key Exchange based on Hyper Elliptic Curve (proposed schem)	۱. توپ ۲. موش ۳. خط ۴. نخ	Hybrid	No need for secure channel, higher speed, and greater efficiency (due to using symmetric encryption for performing encryption and using HECC for key exchange)	O(n)

به‌همین منظور در این مقاله، یک سیستم رمز جدید مبتنی بر DNA برای رمزنگاری ترکیبی متون فارسی و تصاویر پیشنهاد شد؛ به‌طوری‌که در آن، کلید راز اشتراکی با استفاده از تکنیک رمزنگاری HEC تولید شده است. از این‌رو، روش پیشنهادی نیازی به کانال امن برای تبادل کلید راز اشتراکی ندارد. طرح پیشنهادی با کلید رمزنگاری ۶۴۰ بیتی، دارای امنیت رمزنگاری قابل قبول‌تر، فضای کلید رمزنگاری بزرگ‌تر و مقاومت بیشتر در برابر عوامل خارجی است. در این روش، از فناوری‌های مختلف DNA (مانند کدگذاری دیجیتالی DNA، هیبریدسازی DNA و افزودن اطلاعات اضافی به متن رمز شده) استفاده شده است که الگوریتم را امن‌تر می‌سازد. پس از بررسی نتایج حاصل از ارزیابی امنیت و کارایی، می‌توان نتیجه گرفت که الگوریتم پیشنهادی توانسته است تا پارامترهای دقت، سرعت و شکست‌ناپذیری را نسبت به سیستم‌های رمزنگاری موجود بهبود ببخشد. در آینده، سیستم رمز پیشنهادی می‌تواند برای فراهم‌کردن امنیت بلادرنگ در سیستم‌های شبکه‌ای توزیع‌شده استفاده شود. همچنین، می‌توان سیستم پیشنهادی را با موضوعات بسیار پیشرفته‌تری مانند تحقق بسیاری از فناوری‌های امنیتی موجود برای امضا و احراز هویت مورد استفاده قرار داد.

مراجع

[1] A. Bhardwaj, G. V. B. Subrahmanyam, V. Avasthi, and H. Sastry, “Security Algorithms for Cloud Computing,” *Procedia Comput. Sci.*, vol. 85, pp. 535–542, 2016, doi: 10.1016/j.procs.2016.05.215.

[2] M. Mondal and K. S. Ray, “Review on DNA Cryptography,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 2, no. 1, pp. 44–76, 2023, doi: 10.61797/ijbic.v2i1.198.

[3] R. Madhusudhan and R. Shashidhara, "A novel DNA based password authentication system for global roaming in resource-limited mobile environments," *Multimed. Tools Appl.*, vol. 79, no. 3–4, pp. 2185–2212, 2020, doi: 10.1007/s11042-019-08349-8.

[4] B. T. Hammad, A. M. Sagheer, I. T. Ahmed, and N. Jamil, "A comparative review on symmetric and asymmetric dna-based cryptography," *Bull. Electr. Eng. Informatics*, vol. 9, no. 6, pp. 2484–2491, 2020, doi: 10.11591/eei.v9i6.2470.

[5] T. M. Fernandez-Carames, "From Pre-Quantum to Post-Quantum IoT Security: A Survey on Quantum-Resistant Cryptosystems for the Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6457–6480, 2020, doi: 10.1109/JIOT.2019.2958788.

[6] T. Anwar, A. Kumar, and S. Paul, "DNA Cryptography Based on Symmetric Key Exchange," *Int. J. Eng. Technol.*, vol. 7, no. 3, pp. 938–950, 2015.

[7] V. Kolate and R. Joshi, "An Information Security Using DNA Cryptography along with AES Algorithm," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 1, pp. 183–192, 2021, doi: 10.17762/turcomat.v12i1S.1607.

[8] S. Namasudra and G. C. Deka, *Advances of DNA Computing in Cryptography*, 1st ed. New York, NY, USA: Chapman & Hall/CRC, 2018, doi: 10.1201/9781351011419.

[9] N. A. N. Abdullah *et al.*, "A Theoretical Comparative Analysis of Dna Techniques Used in Dna Based Cryptography," *J. Sustain. Sci. Manag.*, vol. 17, no. 5, pp. 165–178, 2022, doi: 10.46754/jssm.2022.05.014.

[10] S. Namasudra and G. C. Deka, *Advances of DNA Computing in Cryptography*, 1st ed. New York, NY, USA: Chapman & Hall/CRC, 2018, doi: 10.1201/9781351011419.

[11] Monika and S. Upadhyaya, "Secure Communication Using DNA Cryptography with Secure Socket Layer (SSL) Protocol in Wireless Sensor Networks," *Procedia Comput. Sci.*, vol. 70, pp. 808–813, 2015, doi: 10.1016/j.procs.2015.10.121.

[12] P. Barman and B. Saha, "An Efficient Hybrid Elliptic Curve Cryptography System with DNA Encoding," *Int. Res. J. Comput. Sci.*, vol. 2, no. May, pp. 33–39, 2015.

[13] S. K. Pujari, G. Bhattacharjee, and S. Bhoi, "A Hybridized Model for Image Encryption through Genetic Algorithm and DNA Sequence," *Procedia Comput. Sci.*, vol. 125, pp. 165–171, 2018, doi: 10.1016/j.procs.2017.12.023.

[14] B. Vikas, A. K. Akshay, S. P. M. Thanneeru, U. M. V. Raghuram, and K. S. Bhargav, "A novel DNA- and PI-based key generating encryption algorithm," *Information Systems Design and Intelligent Applications. Advances in Intelligent Systems and Computing*, vol. 672, pp. 945-954, 2018, doi: 10.1007/978-981-10-7512-4_94.

[15] AA. Sharaieh, A. Edinat, and S. AlFarraji, "An Enhanced Polyalphabetic Algorithm on Vigenerecipher with DNA-Based Cryptography," *IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA)*, Aqaba, Jordan, 2018, pp. 1-6, doi: 10.1109/AICCSA.2018.8612860.

[16] B. D. Patnala and R. Kiran Kumar, "A Novel Level-Based DNA Security Algorithm Using DNA Codons," *Computational Intelligence and Big Data Analytics. SpringerBriefs in Applied Sciences and Technology()*, pp. 1-13, 2019, doi: 10.1007/978-981-13-0544-3_1.

[17] A. H. Saleh, A. S. Yousif, and F. Y. H. Ahmed, "Information Hiding for Text Files by Adopting the Genetic Algorithm and DNA Coding," *IEEE 10th Symposium on Computer Applications & Industrial Electronics (ISCAIE)*, Malaysia, 2020, pp. 220-223, doi: 10.1109/ISCAIE47305.2020.9108842.

[18] S. Balaji, Á. Rocha, and Y.-N. Chung, “Intelligent Communication Technologies and Virtual Mobile Networks Lecture Notes on Data Engineering and Communications Technologies,” *ICICV*, vol. 33, 2019.

[19] M. Irfan Alam and S. N. Singh, “Designing and Implementing Cloud Security Using Multi-layer DNA Cryptography in Python,” *Trends in Wireless Communication and Information Security*, vol. 740, pp. 375-385, 2021, doi: 10.1007/978-981-33-6393-9_38.

[20] P. Prasanna and R. K J, “DNA Based Cryptography with Dual Encryption Using Multiple Cloud,” *Int. J. Res. Appl. Sci. Eng. Technol*., vol. 10, no. 7, pp. 4073–4079, 2022, doi: 10.22214/ijraset.2022.45925.

[21] A. AL-Wattar, “A New Proposed Public Key Cryptography Based on Bio Strands,” *Tech. Rom. J. Appl. Sci. Technol*., vol. 5, pp. 21–28, 2023, doi: 10.47577/technium.v5i.8238.

[22] R. Joshi, M. C. Trivedi, V. Goyal, and D. Bhati, “DNA Sequence in Cryptography: A Study,” *Advances in Data and Information Sciences. Lecture Notes in Networks and Systems*, vol. 522, pp. 557-563, 2023. doi: 10.1007/978-981-19-5292-0_53.

[23] M. Karimi and W. Haider, “Cryptography using DNA Nucleotides,” *Int. J. Comput. Appl*., vol. 168, no. 7, pp. 16–18, 2017, doi: 10.5120/ijca2017914420.

[24] A. Gahlaut, A. Bharti, Y. Dogra, and P. Singh, “DNA based cryptography,” *Commun. Comput. Inf. Sci*., vol. 750, pp. 205–215, 2017, doi: 10.1007/978-981-10-6544-6_20.

[25] A. Rokade, “An Short Introduction to The Data Encryption Algorithm (Des , Aes , Blowfish , Dna Cryptography and Quantum Cryptography),” *International Research Journal of Modernization in Engineering Technology and Science* , vol. 04, no. 10, pp. 997–1000, 2022.

[26] M. A. Iliyasu, O. A. Abisoye, S. A. Bashir, and J. A. Ojeniyi, “A Review of Dna Cryptograhic Approaches, ” *IEEE 2nd International Conference on Cyberspac (CYBER NIGERIA)*, Abuja, Nigeria, 2021, pp. 66-72, doi: 10.1109/CYBERNIGERIA51635.2021.9428855.

[27] S. Pramanik and S. K. Setua, “DNA cryptography,” *7th Int. Conf. Electr. Comput. Eng. ICECE*, 2012, pp. 551–554, doi: 10.1109/ICECE.2012.6471609.

[28] S. Singh and Y. Sharma, “A review on DNA based cryptography for data hiding,” *Proc. Int. Conf. Intell. Sustain. Syst. (ICISS)*, 2019, pp. 282–285, doi: 10.1109/ISS1.2019.8908026.

[29] A. Hazra, S. Ghosh, and S. Jash, “A Review on DNA Based Cryptographic Techniques,” *Int. J. Netw. Secur*., vol. 20, no. 6, pp. 1093–1104, 2018, doi: 10.6633/IJNS.201811_20(6).10.

[30] K. Gupta and S. Singh, “DNA Based Cryptographic Techniques: A Review*,” Int. J. Adv. Res. Comput. Sci. Softw. Eng*., vol. 3, no. 3, 2013.

[31] B. B. Raj and V. Ceronmani Sharmila, “An survey on DNA based cryptography,” *Int. Conf. Emerg. Trends Innov. Eng. Technol. Res. ICETIETR*, 2018, pp. 1–3, doi: 10.1109/ICETIETR.2018.8529075.

JOURNAL OF SOUTHERN COMMUNICATION ENGINEERING
ISLAMIC AZAD UNIVERSITY BUSHEHR BRANCH

E-ISSN: 2980-9231
https://sanad.iau.ir/journal/jce

https://doi.org/10.30495/jce.2025.1993480.1335

Vol. 14/ No. 56/2025

Research Article

Design and Fabrication of a Multiple Input/Multiple Output Aperture-Coupled Multilayer Microstrip Patch Antenna for X-band Applications

Faranak Zarghamian[1] | Mahdi Zavvari*[2]

[1]Department of Electrical Engineering, Urmia branch, Islamic Azad University, Urmia, Iran, f_zarghami88@gmail.com

[2]Antenna and Microwave Research Center, Urmia branch, Islamic Azad University, Urmia, Iran, mahdi.zavvari@iau.ac.ir

Correspondence
Mahdi Zavvari, Associate Professor, Antenna and Microwave Research Center, Urmia branch, Islamic Azad University, Urmia, Iran, mahdi.zavvari@iau.ac.ir

Main Subjects:
Microstrip Patch Antenna Design

Paper History:
Received: 23 April 2024
Revised: 12 July 2024
Accepted: 22 July 2024

Abstract

This paper presents the design and fabrication of a novel multilayer MIMO antenna constructed from two low-cost FR4 substrates and three conductor layers. The lower substrate houses the microstrip feed lines connected to the ports along with the ground plane. The design incorporates λ/4 microstrip impedance matching elements attached to the feed lines, with four rectangular slots etched on the ground plane and positioned vertically above these feed lines. The antenna employs an aperture-coupled feeding technique where signals are guided from the feed lines through the slots to quarter-circle-shaped radiating patches on the upper substrate. The proposed antenna demonstrates multiband operation with three distinct impedance bandwidths covering 7.12-7.54 GHz, 8.27-8.72 GHz, and 10.33-10.96 GHz. The octagonal-shaped antenna structure measures 26mm along its longer side and 10mm along its shorter side. With its stable radiation pattern and acceptable gain characteristics, this compact multilayer design proves suitable for various X-band applications while maintaining a cost-effective fabrication approach using standard FR4 material. The combination of impedance matching elements, slot-coupled feeding, and multilayer architecture contributes to the antenna's effective performance across multiple frequency bands.

Keywords: Microstrip antenna, X-band, Aperture-Coupled, Multi-Input Multi-Output.

Highlights

- A new MIMO antenna with a gate-coupled feeding method was introduced. In this design, the proposed MIMO antenna consists of four microstrip patch antennas side by side.
- The proposed MIMO antenna can cover the frequency bands of 7/54 - 7/12 GHz, 8/72 - 8/27 GHz, and 10/96 - 10/33 GHz in each port.
- In all three frequency bands, the gain of MIMO antenna is positive, and the gain fluctuations are acceptable.

Citation: F. Zarghamian, and M. Zavvari, "Design and Fabrication of a Multiple Input/Multiple Output Aperture-Coupled Multilayer Microstrip Patch Antenna for X-band Applications," *Journal of Southern Communication Engineering*, vol. 14, no. 56, pp. 27-41, 2025, doi:10.30495/jce.2025.1993480.1335 [in Persian].

1. Introduction

The rapid advancement of wireless communication systems has intensified the need for high-performance antennas that can effectively support multiple-input multiple-output (MIMO) technology. MIMO systems provide substantial improvements over conventional single-antenna setups, offering increased channel capacity, better spectral efficiency, and enhanced resilience against multipath fading effects. These advantages are especially valuable for contemporary applications like 5G networks, satellite communications, radar systems, and Internet of Things devices that require dependable high-speed data transfer.

This research concentrates on developing a MIMO antenna specifically optimized for X-band (8–12 GHz) applications, a frequency range extensively employed in satellite communications, military radar systems, and scientific exploration due to its optimal balance between atmospheric penetration capabilities and available bandwidth. However, creating MIMO antennas for this frequency band involves overcoming several design obstacles, such as ensuring strong isolation between antenna elements, preserving compact form factors, and achieving wide bandwidth coverage.

Existing studies have investigated various approaches to tackle these challenges, including aperture-coupled feeding mechanisms, defected ground structures, and multi-layer configurations. Among these, aperture-coupled feeding stands out for its benefits like minimized spurious radiation and enhanced impedance matching characteristics, making it particularly suitable for MIMO antenna implementations. The current study expands upon these established techniques by introducing an innovative multi-layer MIMO antenna design incorporating aperture-coupled feeding, with the goal of attaining exceptional performance metrics regarding bandwidth, port isolation, and radiation efficiency.

The key goals of this investigation include developing a space-efficient, multi-band MIMO antenna with aperture-coupled feeding for X-band applications while simultaneously optimizing the antenna's geometry to achieve superior port isolation and directional radiation characteristics. Through these efforts, this work aims to contribute meaningfully to the progression of MIMO antenna technology, presenting a viable solution for emerging wireless communication system requirements.

2. Antenna Design and Simulation

The proposed MIMO antenna is designed with a multi-layer structure to achieve optimal performance in the X-band. The antenna consists of two FR4 substrates (relative permittivity $\varepsilon_r = 4.4$, loss tangent $\tan \delta = 0.02$) and three conductive layers. The lower substrate houses the microstrip feed lines and a ground plane with four ports, while the upper substrate supports the radiating patches. The ground plane between the two substrates features rectangular slots that facilitate aperture-coupled feeding, enabling efficient energy transfer from the feed lines to the radiating patches.

The radiating patches are designed as quarter circles, strategically placed to minimize mutual coupling and enhance isolation between the MIMO elements. The feed lines are terminated with $\lambda/4$ microstrip stubs to ensure proper impedance matching, which is critical for maximizing power transfer and minimizing reflections. The orthogonal arrangement of the slots and patches further improves isolation by reducing electromagnetic coupling between adjacent elements.

To validate the design, extensive simulations were conducted using ANSYS HFSS, a high-frequency electromagnetic simulation tool. The simulations focused on key performance metrics, including:

- **Impedance matching (S-parameters):** The antenna demonstrates three clear operational bands covering 7.12-7.54 GHz, 8.27-8.72 GHz, and 10.33-10.96 GHz, maintaining S_{11} parameters below -10 dB throughout these ranges, confirming effective power transfer efficiency.
- **Isolation between ports:** Through the orthogonal configuration of slots and patches, the design achieves exceptional port-to-port isolation exceeding 20 dB, as verified by S_{21} and S_{31} parameter measurements for both adjacent and co-linear port combinations.
- **Radiation patterns:** The antenna produces stable directional radiation patterns across all operational bands, reaching a peak gain of 4 dB in the highest frequency band (10.33-10.96 GHz), with pattern consistency that enables reliable beamforming applications.

Parametric optimization studies were conducted to refine key geometric parameters including slot dimensions (length/width) and radiating patch size. The investigation demonstrated a strong correlation between slot length and resonant frequency characteristics, providing valuable tuning capability for application-specific requirements. Simulation results verify the antenna successfully achieves all design targets, delivering a reliable multi-band MIMO solution for X-band communication systems with:

- Frequency-agile performance through slot length adjustment
- Consistent radiation characteristics across operational bands
- Maintained isolation and impedance matching properties
- Compact form factor preservation

The design's tunability and stable performance metrics make it particularly suitable for adaptive X-band systems requiring reconfigurable frequency operation

3. Antenna Fabrication and Experimental Results

The fabricated MIMO antenna prototype was experimentally characterized following its implementation using standard PCB techniques. The manufacturing process involved precise etching of microstrip feed lines and radiating patches on FR4 substrates, followed by careful layer alignment and bonding with 0.1mm positional tolerance. SMA connectors were integrated with all four ports to enable comprehensive testing.

Experimental measurements using a vector network analyzer confirmed the antenna's triple-band operation, showing good agreement with simulated results across the 7.1-7.6 GHz, 8.2-8.8 GHz, and 10.3-11.0 GHz frequency ranges. The measured port isolation consistently exceeded 20dB throughout all operational bands. Far-field measurements revealed peak gains of 0.9dB, 3.1dB, and 3.8dB at center frequencies of 7.4GHz, 8.5GHz, and 10.7GHz, respectively.

The prototype demonstrates several advantages compared to existing designs, including its compact $26\times26\times3.2mm^3$ form factor, simultaneous multi-band operation, and high port isolation. These characteristics, combined with the design's cost-effectiveness and straightforward fabrication process, make it particularly suitable for practical X-band MIMO system implementations. The close correlation between simulated and measured results validates the design methodology and confirms the antenna's real-world performance capabilities.

4. Conclusions

This paper presents the design, simulation, fabrication, and testing of a novel multi-layer MIMO antenna with aperture-coupled feeding for X-band applications. The antenna's innovative geometry features quarter-circular radiating patches and orthogonal slots, enabling high isolation between ports and stable directional radiation patterns. Simulation and experimental results demonstrate triple-band operation at 7.12-7.54 GHz, 8.27-8.72 GHz, and 10.33-10.96 GHz, with the latter two bands falling within the X-band range.

The key achievements of this work include:

- A compact and cost-effective design using FR4 substrates and standard PCB fabrication techniques, ensuring affordability and scalability for mass production.
- High performance with excellent impedance matching, port isolation exceeding 20 dB, and directional radiation characteristics suitable for MIMO applications in satellite communications, radar systems, and IoT devices.
- Experimental validation demonstrating close agreement between simulated and measured results, confirming the design's reliability.

Future work could investigate extended bandwidth capabilities, adaptation for higher frequency ranges, or incorporation of reconfigurable technologies to enable dynamic frequency adjustment. The developed antenna constitutes an important advancement in MIMO antenna technology, providing a viable approach to address contemporary wireless communication requirements.

5. Acknowledgement

The authors acknowledge the Antenna and Microwave Research Center at Islamic Azad University, Urmia branch, for their support of this research. We are grateful to our colleagues for their valuable insights and technical expertise that contributed significantly to this work.

References

[1] A. Ahmad, D.-y. Choi, and S. Ullah, “A compact two elements MIMO antenna for 5G communication,” *Scientific Reports*, vol. 12, p.3608, 2022, doi: 10.1038/s41598-022-07579-5.

[2] T. Aribi, T. Sedghi, and R. Khaje-Mohammadlou, “Multi-Band Compact MIMO Antenna for New Generations of Mobile Applications & IoT,” *J. South. Comm. Eng.,* vol. 12, no. 45, pp. 19-29, 2022, doi: 10.30495/jce.2022.690858.

[3] D. Sharma, B. K. Kanaujia, and S. Kumar, "Compact multi-standard planar MIMO antenna for IoT/WLAN/Sub-6 GHz/X-band applications," *Wireless Networks*, vol. 27, pp. 2671–2689, 2021, doi: 10.1007/s11276-021-02612-3.

[4] S. Rezaee and Y. Zehforoosh, "Design of a Planar Multiband Antenna Using Metamaterials," *J. South. Comm. Eng*., vol. 11, no. 43, pp. 15-26, 2022, doi: 10.30495/jce.2022.689028.

[5] M. Lanza *et al.*, "Memristive technologies for data storage, computation, encryption, and radio-frequency communication," *Science*, vol. 376, no. 6597, 2022, doi: 10.1126/science.abj997.

[6] D. Serghiou, M. Khalily, V. Singh, A. Araghi, and R. Tafazolli, "Sub-6 GHz Dual-Band 8 × 8 MIMO Antenna for 5G Smartphones," in *IEEE Antennas and Wireless Propagation Letters*, vol. 19, no. 9, pp. 1546-1550, Sept. 2020, doi: 10.1109/LAWP.2020.3008962.

[7] Y. Zehforoosh and M. Zavvari, "A novel MIMO antenna with an improved isolation for UWB and multiband applications," *Analog Integrated Circuits and Signal Processing*, vol. 107, pp. 171-179, 2021, doi: 10.1007/s10470-020-01772-0.

Declaration of Competing Interest: Authors do not have conflict of interest. The content of the paper is approved by the authors.

Author Contributions:
Faranak Zarghamian developed the theory and performed the computations and the analytical methods. **Mahdi Zavvari** provided critical feedback and helped shape the research, analysis and contributed to the final version of the manuscript.

E-ISSN: 2980-9231
https://sanad.iau.ir/journal/jce

مهندسی مخابرات جنوب
دانشگاه آزاد اسلامی واحد بوشهر

https://doi.org/10.30495/jce.2025.1993480.1335

مقاله پژوهشی

طراحی و ساخت یک آنتن پچ مایکرواستریپ با تزویج دریچه‌ای به‌صورت چند ورودی/چند خروجی و ساختار چند لایه برای کاربردهای باند X

فرانک ضرغامیان[1] | مهدی زواری*[2]

[1]گروه مهندسی برق، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه، ایران.
f_zarghami88@gmail.com

[2]مرکز تحقیقات آنتن و مایکروویو، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه، ایران.
mahdi.zavvari@iau.ac.ir

نویسنده مسئول
*مهدی زواری، دانشیار مرکز تحقیقات آنتن و مایکروویو، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه، ایران.
mahdi.zavvari@iau.ac.ir

موضوع اصلی:
طراحی آنتن پچ مایکرواستریپ

تاریخچه مقاله:
تاریخ دریافت: ۴ اردیبهشت ۱۴۰۳
تاریخ بازنگری: ۲۲ تیر ۱۴۰۳
تاریخ پذیرش: ۱ مرداد ۱۴۰۳

چکیده:

در این مقاله یک آنتن چند ورودی/چند خروجی با ساختار چند لایه طراحی و ساخت شده است. این آنتن از دو زیر لایه ارزان قیمت FR4 و سه لایه هادی تشکیل شده است. در زیر لایه پایینی خطوط مایکرواستریپ و صفحه زمین متصل به پورت‌های چهار گانه وجود دارند. در این ساختار، خطوط مایکرواستریپ توان را از ورودی‌ها تحویل گرفته و در انتهای خطوط تغذیه با اضافه کردن خطوط مایکرواستریپ $\lambda/4$ جهت تطبیق امپدانسی، این توان به‌صورت تشعشعی ساطع می‌شود. در آن‌سوی زیرلایه، صفحه زمین با چهار شیار مستطیلی شکل دقیقاً بالای خطوط $\lambda/4$ و به‌صورت متعامد قرار گرفته است. توان تشعشعی از طریق این شیارها هدایت شده و به‌صورت تزویج دریچه‌ای به پچ‌های تشعشعی بالایی می‌رسند. این پچ ها به شکل ربع دایره بوده و در لایه بالایی زیر لایه فوقانی قرار دارند. آنتن پیشنهادی دارای سه باند فرکانسی کاری در ۷/۱۲-۷/۵۴ و گیگاهرتز، ۸/۲۷- ۸/۷۲ گیگاهرتز و۱۰/۳۳-۱۰/۹۶ گیگاهرتز هست. این آنتن چند لایه به شکل هشت ضلعی بوده و طول اضلاع بزرگ و کوچک آن به ترتیب ۲۶و ۱۰ میلیمتر بوده و کل ضخامت آن به ۳/۲ میلیمتر می‌رسد. این آنتن دارای الگوی تشعشعی و بهره قابل قبول بوده و مناسب برای کاربردهای جهت دار در باند فرکانسی X هست. آنتن بیشینه بهره در حدود ۴ در باند سوم فرکانسی از خود نشان می‌دهد که در مقایسه با سایر آنتن‌های هم‌اندازه مقدار بزرگتری دارد.

کلید واژه‌ها: آنتن مایکرواستریپ، باند X، تزویج دریچه‌ای، چند ورودی/ چند خروجی

تازه‌های تحقیق:

- یک آنتن MIMO جدید با روش تغذیه تزویج دریچه‌ای معرفی شد. در این طراحی، آنتن MIMO پیشنهادی از چهار آنتن پچ مایکرواستریپی در کنار هم تشکیل شده است.
- آنتن MIMO پیشنهادی در هر پورت خود می‌تواند باندهای فرکانسی ۷/۵۴ - ۷/۱۲ گیگاهرتز، ۸/۷۲ - ۸/۲۷ گیگاهرتز و ۱۰/۹۶ - ۱۰/۳۳ گیگاهرتز را پوشش می‌دهد.
- در هر سه باند فرکانسی، بهره آنتن MIMO مثبت بوده و نوسانات بهره قابل قبول است.

۱-مقدمه

آنتن‌های چند ورودی/چند خروجی به خاطر ویژگی‌هایی از قبیل وزن کم، ابعاد کوچک و استفاده مفید از باند فرکانسی در چندین پورت، قسمت مهمی از مخابرات نوین را تشکیل می‌دهند[۱-۲]. دراین‌بین، طراحی یک آنتن چند ورودی/چند خروجی برای باند فرکانسی ماهواره‌ای X-band بسیار مفید به نظر می‌رسد[۳-۴]. محدوده فرکانسی باند X در اکثر ماهواره‌ها برای کاربردهای نظامی، و نیز کاربردهایی چون اکتشافات علمی، استفاده از منابع بیکران فضا، مخابرات، سنجش از دور، آموزش از دور، مکان‌یابی و ناوبری به‌طور قابل‌توجهی گسترش یافته‌اند[۵]. سیستم‌های چند ورودی/چند خروجی در مقایسه با سیستم‌های معمول تک ورودی/تک خروجی، می‌توانند یک ارتباط بیسیم با کارایی طیفی و قابلیت اطمینان بالا را فراهم کنند. سیستم‌های چند ورودی/چند خروجی، جهت ارتقاء کیفیت ارتباط بیسیم در محیط‌های چند مسیری برای سیگنال از چند آنتن به‌عنوان گیرنده و فرستنده استفاده می‌کنند[۶]. با توجه به امتیازات این سیستم‌ها، استفاده از آنها در استانداردهای مختلف فرکانسی از قبیل LTE، WLAN و... همواره توصیه می‌شود. وجود چندین مسیر مخابراتی بین فرستنده و گیرنده ضریب اطمینان سیستم را به نحو چشمگیری افزایش می‌دهد. به این صورت که در صورت مساعد نبودن یکی از این مسیرها به هر دلیلی، مسیرهای مخابراتی دیگر امکان برقراری ارتباط رادیویی را برای فرستنده و گیرنده فراهم می‌کنند[۷]. سیستم چند ورودی چند خروجی (MIMO) یک راه حل مهم در بهبود ظرفیت کانال برای یک ایستگاه پایه داخلی است. آنتن‌های MIMO در محیط داخلی از چند پارامتر مانند قطبش، پترن تشعشعی و پیکربندی آرایه برای بهبود ظرفیت کانال استفاده می‌کنند. روش چندگانگی قطبش می‌تواند ظرفیت کانال را افزایش دهد ، برای مثال، آنتن با قطبش متفاوت از ظرفیت بیشتر نسبت به آنتن با قطبش یکسان، برخوردار است. نتیجه‌ی مشابهی نیز به دست آمد که در آن سیستم قطبش ترکیبی، ظرفیت بیشتری نسبت به سیستم تک قطبشی ارائه می‌دهد [۷]. در طراحی آنتن MIMO ، ساختار معمولاً به‌صورت آرایه‌ای است. ابعاد کم، اندازه کوچک در طراحی باید برای حفظ کم‌هزینه بودن آنتن با نصب ساده در نظر گرفته شود. اما، ساخت آرایه‌ای با اندازه کوچک، کار ساده‌ای نیست. تزویج متقابل بین عناصر آنتن باید تا حد ممکن کم باشد و ضریب همبستگی کوچک بین آرایه‌ها حفظ شود.

از طرفی روش‌های بسیار متعددی برای تغذیه آنتن‌های مسطح وجود دارد که می‌توانند نسبت به هم مزایا و معایبی داشته باشند. این روش‌های تغذیه به‌طورکلی در دو گروه عمده دسته‌بندی می‌شوند، تماسی و غیر تماسی. در روش‌های تماسی پچ آنتن مستقیماً به منبع موردنظر متصل می‌گردد ولی در روش غیر تماسی انتقال توان از منبع تغذیه به پچ آنتن از طریق میدان‌های مغناطیسی ترویج شده منتقل می‌گردد. به‌این‌ترتیب المان تشعشعی می‌تواند به‌صورت جداگانه بر روی ساختار تغذیه آنتن قرار گرفته و تعویض و تغییر آن نیز میسر می‌شود [۸]. در روش تزویج دریچه‌ای، پچ آنتن طوری قرار گرفته که دو زیر لایه از هم جدا بوده و خط مایکرواستریپ در زیر لایه تحتانی قرار گرفته است، انرژی از طرف خط تغذیه به پچ تشعشعی در روی زیر لایه بالایی القا می‌گردد. در این روش تزویج از طریق یک شکاف که بر روی زمین قرار دارد انجام می‌شود. در این ساختار دو زیر لایه روی هم قرار گرفته و خط تغذیه در لایه زیرین زیر لایه تحتانی قرار دارد. همچنین پچ در لایه بالایی زیر لایه فوقانی قرارگرفته و صفحه زمین بین این دو قرار می‌گیرد[۹].

در سالهای اخیر مطالعات فراوانی در این حوزه انجام گرفته و مقالات متعددی گزارش شده‌اند که از آن جمله لین و همکاران در ۲۰۱۵ یک آنتن چند ورودی/چند خروجی که از یک حلقه رزوناتور دی الکتریک استفاده می‌کند ارائه کردند[۱۰]. در این طراحی از تشعشع کننده های هیبریدی خود مکمل جهت تشکیل آنتن MIMO استفاده شده است. یک خط تغذیه مایکرواستریپ جهت تحریک شکاف پله‌دار و DRA به کار رفته است. با ترکیب رزونانس‌های شکاف پله‌ای و DAR پهنای باند امپدانسی وسیعی به‌دست‌آمده است. آنتن MIMO این مرجع دارای دریچه‌ای به ابعاد ۵۰ × ۳۱ میلیمتر مربع هست. عملکرد این آنتن به‌صورت زیر هست: پهنای باند پورت یک برابر با ۸۲/۷۵٪ و پورت دوم برابر با ۸۵/۲۱٪ ، ایزولاسیون بین پورت‌ها بیشتر از ۲۰ دسی‌بل و الگوی تشعشعی پایدار در باند فرکانسی حاصل شده است.

کین و همکاران در ۲۰۱۲ یک آنتن با تغذیه تزویج دریچه‌ای ارائه کرده‌اند. این آنتن با به‌کارگیری ساختار راه‌راه برای صفحه زمین، عملکرد آنتن را بهبود بخشیده است. این طراحی از دو زیر لایه و یک لایه زمین در بین آن‌ها تشکیل شده است. در این ساختار، خط تغذیه در زیر لایه پایین و پچ تشعشعی در زیر لایه بالا قرار گرفته است. با اعمال یک لایه زمین شیاردار در بین این دو توانسته‌اند به‌طور وسیعی عملکرد آنتن را بهبود ببخشند. به‌طوری‌که بهره آنتن نسبت به حالت زمین ساده، ۳ / ۵ دسی‌بل

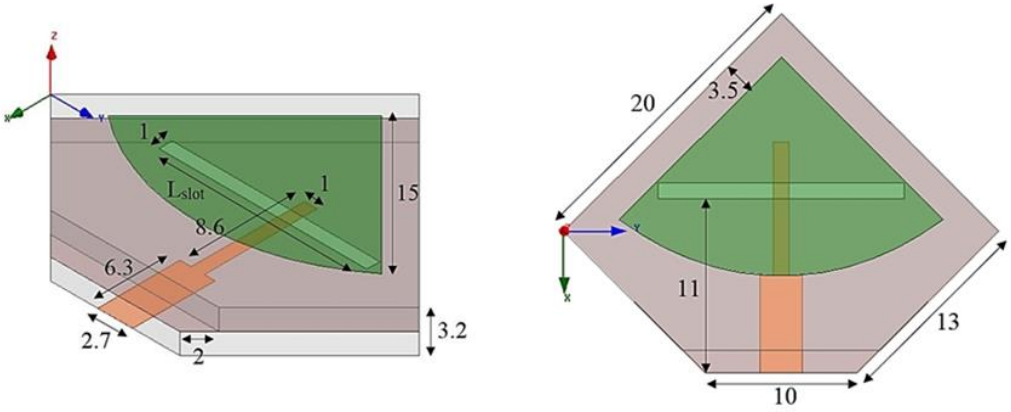
افزایش پیداکرده است. همچنین الگوی تشعشعی آنتن نسبت به حالت با زمین ساده بهتر شده و پهنای بیم نصف توان آن بهبود پیدا کرده است[۱۱].

در طرحی دیگر، ماجد و همکاران در ۲۰۱۴ از المان DRA جهت بهبود پارامترهای آنتن استفاده شده است. به این صورت که یک خط تغذیه در لایه تحتانی زیر لایه و یک شکاف در لایه فوقانی زیر لایه بر روی صفحه زمین، وظیفه انتقال توان تشعشعی از خط مایکرواستریپ را به DRA به عهده دارند[۱۲].

در این مقاله سعی بر آن است که، یک آنتن چند ورودی/چند خروجی با ساختار چند لایه طراحی شود که نوع تغذیه آن به‌صورت تزویج دریچه‌ای بوده و قسمتی از باند فرکانسی X-band در محدوده فرکانسی ۱۲-۸ گیگاهرتز را پوشش دهد. این آنتن با استفاده از تکنولوژی چاپ بر روی فیبر مدار چاپی ساخت و در آزمایشگاه آنتن مورد تست قرار خواهد گرفت.

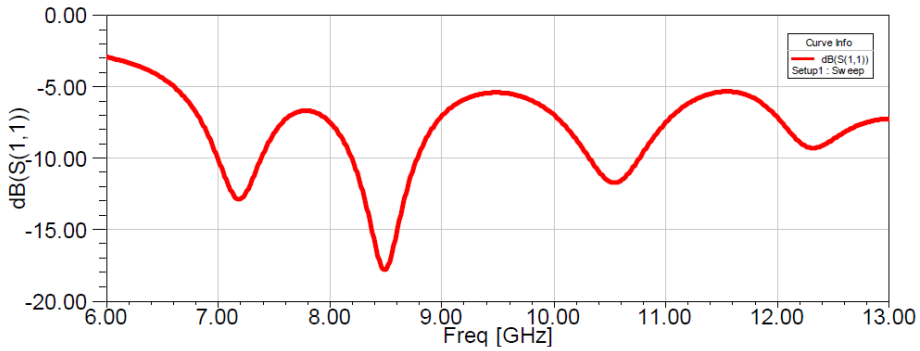
۲- طراحی آنتن پیشنهادی و نتایج

در طرح پیشنهادی این آنتن از دو زیر لایه FR4 و سه لایه هادی تشکیل شده است. در زیر لایه پایینی خطوط مایکرواستریپ و صفحه زمین متصل به پورت‌های چهار گانه وجود دارند. در این ساختار، خطوط مایکرواستریپ توان را از ورودی‌ها تحویل گرفته و در انتهای خطوط تغذیه با اضافه کردن خطوط مایکرواستریپ $\lambda/4$ جهت تطبیق امپدانسی، این توان به‌صورت تشعشعی ساطع می‌شود. در آن‌سوی زیرلایه، صفحه زمین با چهار شیار مستطیلی شکل دقیقاً بالای خطوط $\lambda/4$ و به‌صورت متعامد قرار گرفته است. توان تشعشعی از طریق این شیارها هدایت شده و به‌صورت تزویج دریچه‌ای به پچ های تشعشعی بالایی می‌رسند. این پچ ها به شکل ربع دایره بوده و در لایه بالایی زیر لایه فوقانی قرار دارند. آنتن پیشنهادی دارای سه باند فرکانسی کاری در ۷/۷-۷۱/۲۵ گیگاهرتز، ۸/۹۶-۸/۰۳ گیگاهرتز و ۱۱/۱۲-۱۰/۶۷ گیگاهرتز هست. این آنتن چند لایه به شکل هشت ضلعی بوده و طول اضلاع بزرگ و کوچک آن به ترتیب ۲۶ و ۱۰ میلیمتر بوده و کل ضخامت آن به ۳/۲ میلیمتر می‌رسد. انتظار می‌رود این آنتن دارای الگوی تشعشعی و بهره قابل قبول بوده و مناسب برای کاربردهای جهت دار در باند فرکانسی X هست. آنتن تک المانه در این طراحی مطابق شکل ۱ از یک پچ دایروی به شعاع ۱۵ میلیمتر بر روی یک زیر لایه FR4 با ضخامت ۱/۶ میلیمتر به‌عنوان تشعشع کننده اصلی تشکیل شده است. این زیر لایه دارای ضریب گذردهی نسبی ۴/۴ و تانژانت تلفات ۰/۰۲ هست. در زیر این زیرلایه، زیرلایه دیگری با همان مشخصات وجود دارد که شامل خط تغذیه در لایه پایینی و صفحه زمین در لایه فوقانی هست. به‌این‌ترتیب صفحه زمین بین پچ تشعشعی و خط تغذیه قرار می‌گیرد. در این طراحی جهت هدایت امواج الکترومغناطیسی به پچ تشعشعی یک شکاف مستطیلی شکل در صفحه زمین و درست زیر پچ تشعشعی به‌صورت عمود بر خط تغذیه قرار داده‌ایم. بنابراین مجموعه خط تغذیه و شکاف موجود در صفحه زمین تغذیه تزویج دریچه‌ای را در این آنتن تشکیل می‌دهند. همچنین جهت انجام تطبیق امپدانسی بین خط تغذیه و پورت ورودی، یک خط مایکرواستریپ به طول $\lambda/4$ به انتهای خط تغذیه متصل شده است. در این آنتن به دلیل استفاده از زیر لایه FR4 به ضخامت ۱/۶ میلیمتر، از خط تغذیه



شکل ۱: نمای سه‌بعدی و از بالا از آنتن تک المانه

Figure 1. 3D and top view of proposed sigle element antenna



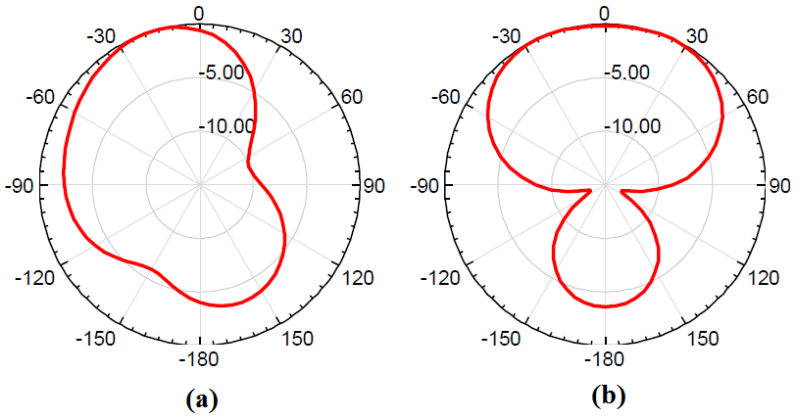
شکل ۲: باند فرکانسی کاری آنتن تک المانه پیشنهادی

Figure 2. Calculated return loss of proposed sigle element antenna

مایکرواستریپ به پهنای ۲/۷ میلیمتر و در ازای ۶/۳ میلیمتر استفاده شده است. در شکل ۱ پیکربندی آنتن تک المانه پیشنهادی نشان داده شده است. تمام ابعاد بر روی شکل نوشته شده و در مقیاس میلیمتر می‌باشند. با توجه به شکل، جهت دسترسی کانکتور SMA به صفحه زمین، زیر لایه فوقانی، مقداری کوچک‌تر از زیرلایه تحتانی انتخاب شده است. در شکل ۲، نتایج حاصل از شبیه‌سازی این آنتن نشان داده شده است. با توجه به شکل، آنتن تک المانه دارای باندهای فرکانسی ۷/۰۲-۷/۳۷ گیگاهرتز، ۸/۱۸-۸/۷۸ گیگاهرتز و۱۰/۳۲-۱۰/۷۶ گیگاهرتز هست. از بین این سه باند فرکانسی، باندهای دوم و سوم در محدوده باند فرکانسی X قرار دارند. شکل ۳ الگوی تشعشعی آنتن تک المانه را در دو صفحه xy و yz به‌صورت مجزا نشان می‌دهد. طبق این شکل‌ها، شاهد یک الگوی تشعشعی دایرکشنال از آنتن در جهت محور Z ها هستیم. در طراحی ارائه شده، طول شکاف زمین یا همان دریچه، نقش مهمی در تعیین فرکانس کاری آنتن ایفا می‌کند.

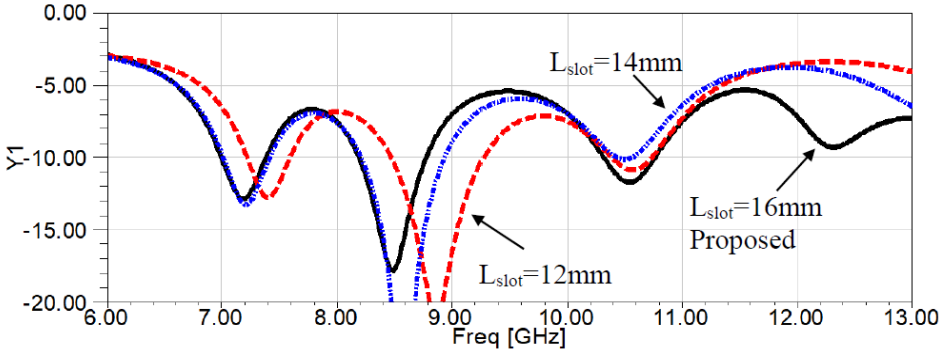
در شکل ۴ منحنی تغییرات S11 آنتن را به ازای تغییرات طول شکاف رسم کرده‌ایم. با توجه به شکل، مشاهده می‌شود که با افزایش طول شکاف، فرکانس کاری آنتن به سمت فرکانس‌های پایین سوق پیدا می‌کند.

آنتن تک المانه در باندهای فرکانسی خود، بیشینه بهره‌هایی معادل با ۰/۰۸۵ دسی‌بل، ۱/۲۷دسی‌بل و ۱/۴۵ دسی‌بل به ترتیب در باندهای فرکانسی اول، دوم و سوم خود به دست آورده است.

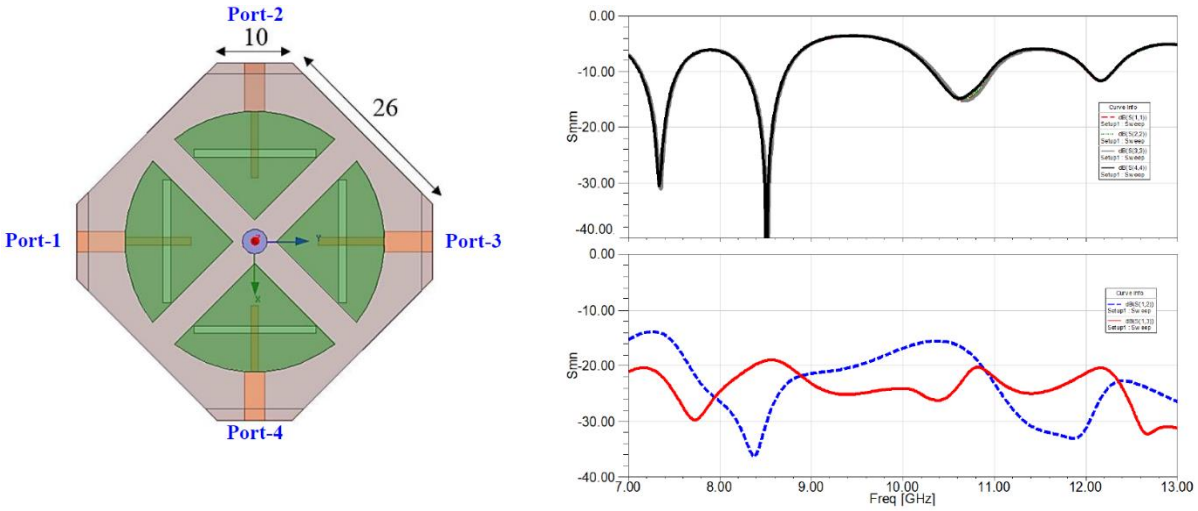


شکل ۳: الگوی تشعشعی آنتن تک المانه در صفحه xy (a) و yz (b)

Figure 3. Radiation pattern of proposed sigle element antenna in (a) xy-plane, (b) yz-plane



شکل ۴: منحنی S11 آنتن تک المانه بر حسب مقادیر مختلف طول شکاف
Figure 4. S11 parameter of proposed sigle element antenna vs. different slot lengths



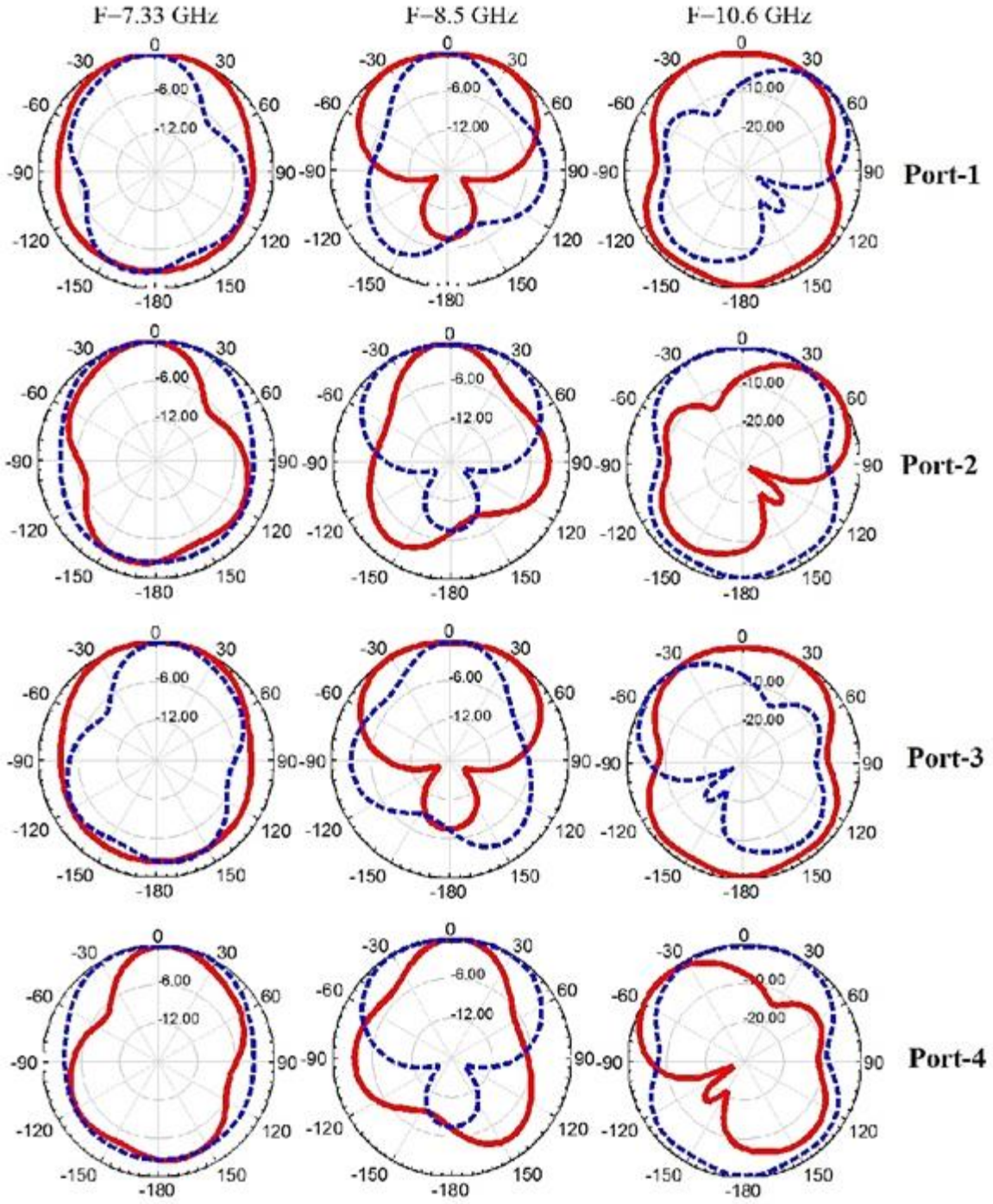
شکل ۵: طرح نهایی آنتن پیشنهادی، نمودار تلفات بازگشتی و ایزولاسیون بین پورت‌های متعامد و هم‌راستا
Figure 5. Final antenna design, and its return loss and isolation between orthogonal and unidirection ports

در ادامه به طرح نهایی آنتن با استفاده از المان‌های طراحی شده می‌پردازیم. شکل ۵ ساختار آنتن چند ورودی/چند خروجی ارائه شده در این مقاله را نشان می‌دهد. به زبان ساده با کنار هم گذاشتن، چهار عدد از آنتن‌هایی که در بخش قبل مورد بررسی قرار دادیم، آنتن چند ورودی/چند خروجی پیشنهادی شکل می‌گیرد. این آنتن از دو زیر لایه FR4 و سه لایه هادی تشکیل شده است. در زیر لایه پایینی خطوط مایکرواستریپ و صفحه زمین متصل به پورت‌های چهار گانه وجود دارند. در این ساختار، خطوط مایکرواستریپ توان را از ورودی‌ها تحویل گرفته و در انتهای خطوط تغذیه با اضافه کردن خطوط مایکرواستریپ $\lambda/4$ جهت تطبیق امپدانسی، این توان به‌صورت تشعشعی ساطع می‌شود. در آن‌سوی زیرلایه، صفحه زمین با چهار شیار مستطیلی شکل به طول L, slot دقیقاً بالای خطوط $\lambda/4$ و به‌صورت متعامد قرار گرفته است.توان تشعشعی از طریق این شیارها هدایت شده و به‌صورت تزویج دریچه‌ای به پچ های تشعشعی بالایی می‌رسند .این پچ ها به شکل ربع دایره بوده و در لایه بالایی زیر لایه فوقانی قرار دارند. طراحی ارائه شده دو زیر لایه از لحاظ جنس و ابعاد مشابه هم هستند، فقط در محل پورت‌های آنتن، جهت دستیابی به صفحه زمین برای لحیم کاری، طول زیر لایه به اندازه ۲ میلیمتر کوتاه‌تر از زیر لایه دیگر انتخاب شده است .همچنین جهت قرار گرفتن و ثابت ماندن دو زیر لایه بر روی هم، با انجام سوراخ‌کاری در وسط آنتن، به کمک یک عدد پیچ و مهره فلزی این کار را انجام داده‌ایم. لازم به ذکر است که هر دوی این موارد در شبیه‌سازی لحاظ شده‌اند، و تأثیر آنها بر روی عملکرد آنتن مورد بررسی قرار گرفته است. این تأثیرگذاری بر روی رفتار آنتن بسیار ناچیز بوده و با تقریب بسیار خوب می‌توان از اثر آنها صرف‌نظر کرد.

برای آنتن پیشنهادی دو پارامتر اسکالر توان بازگشتی از پورت‌ها و ایزولاسیون بین پورت‌ها مطرح هست. با توجه به ساختار آنتن، یک شکل کاملاً متقارن به چشم می‌خورد. لذا انتظار می‌رود که نمودار توان بازگشتی از هر پورت مشابه پورت‌های دیگر باشد یعنی پهنای باند امپدانسی تمام پورت‌ها یکسان خواهد بود. از لحاظ ایزولاسیون، در آنتن پیشنهادی می‌توان دو نوع ایزولاسیون مختلف برای پورت‌ها تعریف کرد. با توجه به ساختار آنتن، یا دو پورت نسبت به هم عمودند (پورت‌های ۱ و ۲ یا ۱ و ۴ یا ۳ و ۲ یا ۳ و ۴) در راستای یکدیگر قرار دارند (پورت‌های ۱ و ۳ یا ۲ و ۴). بنابراین برای بررسی ایزولاسیون یک مورد از هر دو حالت را در نظر خواهیم گرفت. شکل ۵ نمودار پهنای باند امپدانسی آنتن MIMO پیشنهادی را در هر چهار پورت نشان می‌دهد. همانطورکه قبلا گفته شد، هر چهار نمودار یکسان بوده و پهنای باندهای برابری را برای آنتن نمایش می‌دهند. طبق نتایج شبیه‌سازی شده با نرم افزار HFSS Ver.13 ، آنتن MIMO پیشنهادی در هر پورت خود می‌تواند باندهای فرکانسی ۷/۱۲-۷/۵۴ گیگاهرتز، ۸/۲۷-۸/۷۲ گیگاهرتز و۱۰/۳۳-۱۰/۹۶ گیگاهرتز را پوشش دهد. همچنین ایزولاسیون بین پورت‌های آنتن MIMO پیشنهادی در شکل ۵ نمایش داده شده است. این ایزولاسیون شامل، پورت‌های متعامد (پورت‌های ۱ و ۲) و پورت‌های هم‌راستای آنتن (پورت‌های ۱ و ۳) می‌شود. با توجه به شکل مشاهده می‌شود که در بعضی از باندهای فرکانسی پورت‌های متعامد از ایزولاسیون بهتری برخوردارند، و در بعضی از باندهای فرکانسی، پورت‌های هم‌راستا ایزولاسیون قوی‌تری دارند. در ادامه به بررسی پارامترهای برداری آنتن پیشنهادی می‌پردازیم. در شکل ۶ الگوی تشعشعی آنتن MIMO پیشنهادی را در هر سه باند فرکانسی برای پورت‌های آنتن رسم کرده‌ایم. در این شکل الگوی تشعشعی صفحه zx را با رنگ قرمز و الگوی تشعشعی صفحه zy را با رنگ آبی رسم کرده‌ایم. تقریباً می‌توان گفت که آنتن پیشنهادی دارای چندگانگی الگوی تشعشعی ۱ بوده و این قضیه در فرکانس‌های بالا (باند فرکانسی۱۰/۳۳-۱۰/۹۶ گیگاهرتز) بیشتر به چشم می‌خورد. این خاصیت یک ویژگی مهم برای آنتن به‌حساب می‌آید چرا که آنتن MIMO با داشتن الگوهای تشعشعی مجزا برای هر پورت علاوه بر افزایش ایزولاسیون پورت‌ها می‌تواند به‌عنوان آنتنی که قابلیت اسکن فضا را دارد به شمار رود.

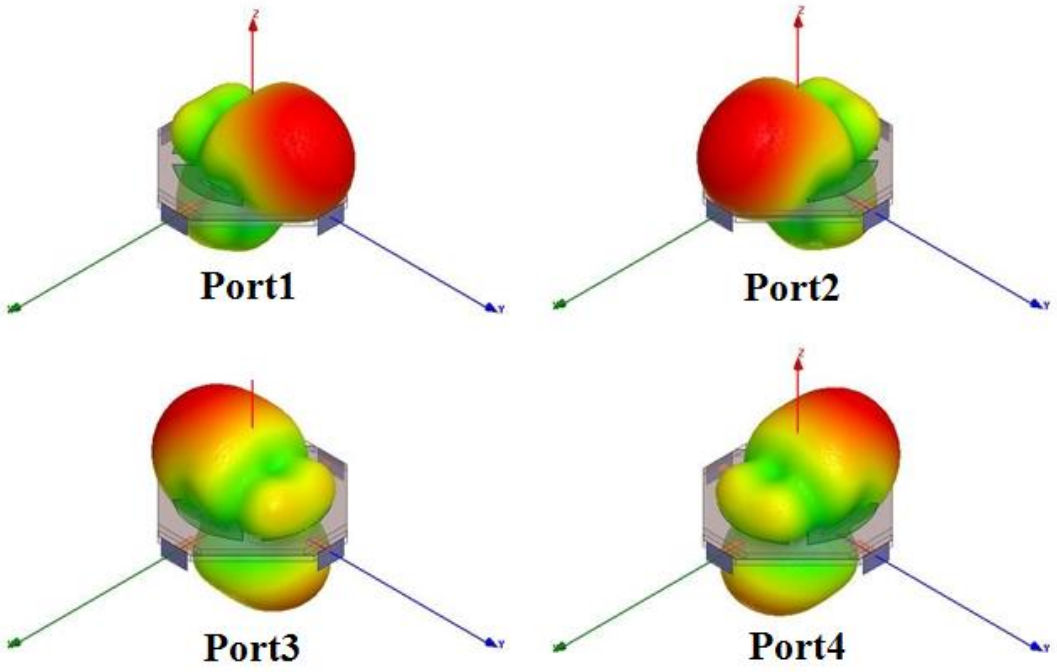
درواقع آنتن می‌تواند با فعال کردن هر کدام از پورت‌های خود فقط قسمتی از فضا را اسکن کرده و به‌این‌ترتیب با آنتن پیشنهادی می‌توان، فضا را به چهار منطقه مجزا تقسیم کرد. ویژگی اسکن فضا به‌عنوان یک پارامتر مهم در رادارهای نظامی و هواشناسی محسوب می‌شود، چرا که به‌جای چرخش مکانیکی آنتن در جهات مختلف، کافی است تا با سوئیچ پورت‌ها این کار به‌صورت الکتریکی انجام شود. برای روشن شدن این موضوع، الگوی تشعشعی سه‌بعدی آنتن پیشنهادی را به ازای فعال کردن پورت‌های چهارگانه مورد تحلیل و بررسی قرار می‌دهیم. شکل ۷ الگوی تشعشعی سه‌بعدی آنتن MIMO پیشنهادی را در فرکانس۱۰/۶ گیگاهرتز برای پورت‌های آنتن به‌صورت جداگانه نمایش می‌دهد. همان‌طور که در شکل مشاهده می‌شود، در هنگام تحریک هر پورتی، الگوی تشعشعی آن درست در جهت مقابل آن تشکیل می‌شود. این موضوع می‌تواند به دلیل حضور پورت‌ها و پچ های تشعشعی دیگر در اطراف آن باشد که مانند یک منعکس‌کننده به الگوی تشعشعی آنتن جهت می‌دهند. بنابراین با توجه به ساختار آنتن که در شکل ۵ نمایش داده شده است، انتظار می‌رود که الگوی تشعشعی مربوط به پورت یک متمایل در جهت محور $+Y$ ، الگوی تشعشعی پورت دو متمایل در جهت محور $+X$ ، الگوی تشعشعی پورت سه متمایل در جهت محور $-Y$ و الگوی تشعشعی پورت چهار متمایل در جهت محور $-X$ باشند.

در شکل ۸ نمودار شبیه‌سازی شده بهره این آنتن نشان داده شده است. همان‌طور که در شکل دیده می‌شود در هر سه باند فرکانسی (مستطیل‌های آبی‌رنگ)، بهره آنتن مثبت بوده و نوسانات بهره قابل قبول است. در باند 3دسی‌بل و در باند فرکانسی اول ، آنتن دارای بیشینه بهره ۰/۹ دسی‌بل، در باند فرکانسی دوم ۳/۱ دسی‌بل و در باند فرکانسی سوم ۳/۸ دسی‌بل هست.

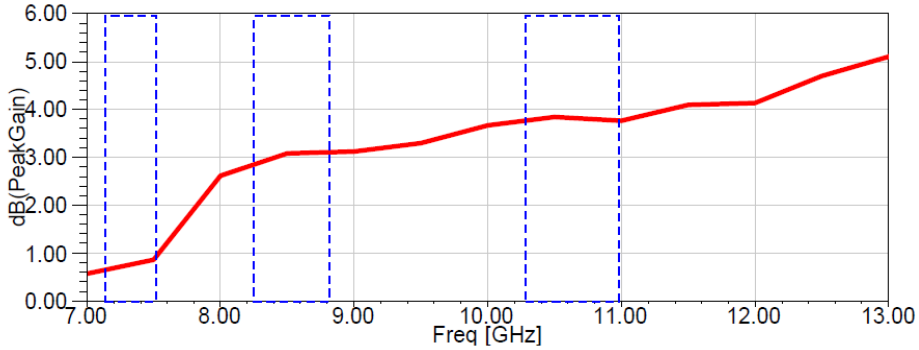


شکل۶: الگوی تشعشعی آنتن MIMO پیشنهادی را در هر سه باند فرکانسی برای پورت‌های آنتن

Figure 6. Radiation pattern of final antenna design at three frequency bands and for different ports



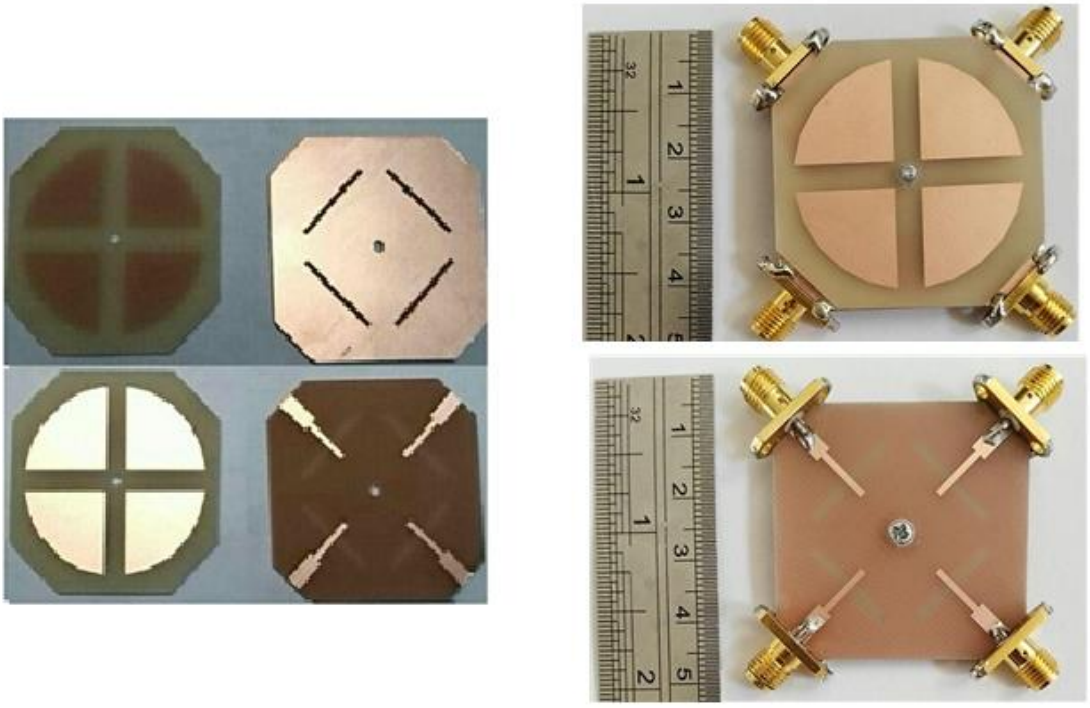
شکل۷: الگوی تشعشعی سه‌بعدی آنتن پیشنهادی
Figure 7. 3D radiation pattern of final antenna design



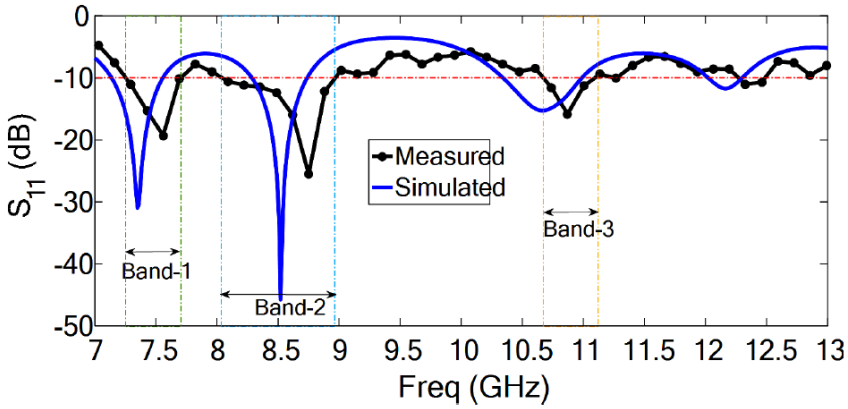
شکل ۸: نمودار محاسبه شده از نتایج شبیه‌سازی بهره آنتن
Figure 8. Calculated gain of final antenna design obtained from simulation

۳- ساخت و تست آنتن پیشنهادی

آنتن پیشنهادی پس از تحلیل و بررسی پارامترهای مهم آن، با تکنیک چاپ بر روی فیبر SMA مدار چاپی در دو لایه ساخته شد و در نهایت توسط یک پیچ به هم متصل شد. سپس چهار عدد کانکتور SMA ، دو پیچ کوتاه از نوع مادگی جهت ارتباط با آنتن به پورت‌های آن لحیم کاری شد. در شکل ۹ طرح نهایی ساخته شده قبل و بعد از مونتاژ مشاهده می‌شود. آنتن پیشنهادی در این کار بعد از ساخت در آزمایشگاه آنتن دانشگاه آزاد اسلامی واحد ارومیه توسط دستگاه تحلیلگر شبکه مورد تست و اندازه‌گیری قرار گرفت. شکل ۱۰ نمودار افت بازگشتی آنتن MIMO پیشنهادی را نشان می‌دهد. با توجه به شکل، همخوانی خوبی بین نمودارهای تست و شبیه‌سازی به چشم می‌خورد. طبق نتایج تست، آنتن پیشنهادی دارای باندهای فرکانسی به ترتیب، باند فرکانسی اول۷/۲۵-۷/۷۱ گیگاهرتز، باند فرکانسی دوم۸/۰۳-۸/۹۶ گیگاهرتز و باند فرکانسی سوم۱۰/۶۷-۱۱/۱۲ گیگاهرتز هست. همانند حالت شبیه‌سازی، باندهای فرکانسی دوم و سوم در محدوده باند فرکانسی X قرار دارند و می‌توانند در این باند فرکانسی فعالیت نمایند. همچنین با توجه به الگوی تشعشعی و بهره قابل قبول آنتن پیشنهادی، این آنتن را می‌توان به‌عنوان یک کاندید خوب برای به‌کارگیری در باند فرکانسی X برای کاربردهای راداری در نظر گرفت.



شکل ۹: لایه‌های مختلف ساخته شده آنتن و شکل نهایی آن
Figure 9. Fabricated layers and final assembled antenna structure



شکل ۱۰: نمودار تست و شبیه‌سازی شده افت بازگشتی آنتن MIMO پیشنهادی در پورت یک
Figure 10. Simulated and measured return loss for proposed antenna

جدول ۱: مقایسه نتایج مقاله با کارهای سایرین
Table 1. Comparison of charcteristics of this work and others

Ref.	Operation band (GHz)	Isolation(dB)	Gain	Dimenssions(mm)
[13]	6.6-7.6&8.3-10	17	1.7	**42*17**
[14]	5.55-5.86&7.93-8.36	18.7	3.5	**21.29*50.54**
[15]	3.3-3.6&5-6&7.2-8.4	15	5	**58*45**
This work	7.02-7.37&7.78-8.18&10.32-10.76	40	4	**26*26**

۴-نتیجه‌گیری

در این کار یک آنتن MIMO جدید با روش تغذیه تزویج دریچه‌ای معرفی شد که از چهار آنتن پچ مایکرواستریپی در کنار هم تشکیل شده است. آنتن‌های تک المانه بکار رفته در ساختار نهایی دارای باندهای فرکانسی ۷/۰۲-۷/۳۷ گیگاهرتز، ۸/۱۸-۷/۷۸ گیگاهرتز و ۱۰/۳۲-۱۰/۷۶ گیگاهرتز هست که از بین این سه باند فرکانسی، باندهای دوم و سوم در محدوده باند فرکانسی X قرار دارند. آنتن تک المانه دارای الگوی تشعشعی و بهره قابل قبولی هست. نتایج شبیه‌سازی نشان می‌دهد آنتن MIMO

پیشنهادی در هر پورت خود می‌تواند باندهای فرکانسی ۷/۱۲-۷/۵۴ گیگاهرتز، ۸/۲۷-۸/۷۲ گیگاهرتز و۱۰/۳۳-۱۰/۹۶ گیگاهرتز را پوشش دهد و در هر سه باند فرکانسی، بهره آنتن MIMO مثبت بوده و نوسانات بهره قابل قبول است و نیز آنتن پیشنهادی دارای چندگانگی الگوی تشعشعی هست. نتایج ساخت و تست آنتن همخوانی خوب نتایج شبیه‌سازی و عملی را نشان می‌دهد.

مراجع

[1] A. Ahmad, D.-y. Choi, and S. Ullah, “A compact two elements MIMO antenna for 5G communication,” *Scientific Reports*, vol. 12, p. 3608, 2022, doi: 10.1038/s41598-022-07579-5.

[2] T. Aribi, T. Sedghi, and R. Khaje-Mohammadlou, “Multi-Band Compact MIMO Antenna for New Generations of Mobile Applications & IoT,” *J. South. Comm. Eng.,* Vol. 12, pp. 19-29, 2022, doi: 10.30495/jce.2022.690858.

[3] D. Sharma, B. K. Kanaujia, and S. Kumar, "Compact multi-standard planar MIMO antenna for IoT/WLAN/Sub-6 GHz/X-band applications," *Wireless Networks*, vol. 27, pp. 2671–2689, 2021, doi: 10.1007/s11276-021-02612-3.

[4] S. Rezaee and Y. Zehforoosh, "Design of a Planar Multiband Antenna Using Metamaterials," *J. South. Comm. Eng*., vol. 11, pp. 15-26, 2022, doi: 10.30495/jce.2022.689028.

[5] M. Lanza *et al.,* "Memristive technologies for data storage, computation, encryption, and radio-frequency communication," *Science*, vol. 376, p. 6597, 2022, doi: 10.1126/science.abj9979.

[6] D. Serghiou *et al.,* "Sub-6 GHz Dual-Band 8 × 8 MIMO Antenna for 5G Smartphones," *IEEE Antennas and Wireless Propagation Letters*, vol. 19, pp. 1546-1550, 2020, **doi:** 10.1109/LAWP.2020.3008962.

[7] Y. Zehforoosh and M. Zavvari, "A novel MIMO antenna with an improved isolation for UWB and multiband applications," *Analog Integrated Circuits and Signal Processing*, vol. 107, pp. 171-179, 2021, doi: 10.1007/s10470-020-01772-0.

[8] M. H. Ashouri and S. Fakhte, "Reduction of Mutual Coupling Between Two Circularly Polarized Magneto-Electric Dipole Antennas Using Metasurface Wall Polarization Converter for 5G Application," *Majlesi J. Elect. Eng.*, vol. 18, pp. 55-63, 2024, doi: 10.30486/mjee.2023.1991539.1180.

[9] Y.-M. Zhang and S, Zhang, "A Novel Aperture-Loaded Decoupling Concept for Patch Antenna Arrays," *IEEE Transactions on Microwave Theory and Techniques*, vol. 69, pp. 4272-4283, 2021, doi**:** 10.1109/TMTT.2021.3085904.

[10] M. Lin and Z. Li, “A Compact 4 x 4 Dual Band-notched UWB MIMO Antenna with High Isolation,” *IEEE 6th International Symposium on Microwave*, Antenna, Propagation, and EMC Technologies (MAPE), pp. 126-128, 28-30 Oct. 2015, **doi:** 10.1109/MAPE.2015.7510281.

[11] K. Qin, M. Li, H. Xia, and J. Wang, "A New Compact Aperture-Coupled Microstrip Antenna with Corrugated Ground Plane," *IEEE Anten. Wireless Propag. Lett.*, vol. 11, pp. 807-810, 2012, **doi:** 10.1109/LAWP.2012.2208212.

[12] A. H. Majeed, A. S. Abdullah, F. Elmegri, K. H. Sayidmarie, R. A. Abd-Alhameed, and J. M. Noras, "Aperture-Coupled Asymmetric Dielectric Resonators Antenna for Wideband Applications," *IEEE Anten. Wireless Propag. Lett.*, vol. 13, pp. 927-930, 2014, **doi:** 10.1109/LAWP.2014.2322779.

[13] N. Pouyanfar, C. Ghobadi, and J. Nourinia, “A compact multi-band MIMO antenna with high isolation for C and X bands using defected ground structure,” *Radioengineering*, vol. 27, pp.686–693, 2018, doi: 10.13164/re.2018.0686.

[14] V. Satam and S. Nema, "Defected ground structure planar dual element MIMO antenna for wireless and short range RADAR application," *IEEE International Conference on Signal Processing, Informatics,*

Communication and Energy Systems (SPICES), Kozhikode, India, 2015, pp. 1-5, doi: 10.1109/SPICES.2015.7091441.

[15] N. Jaglan *et al.,* "Triple band notched DG-CEBG structure based UWB MIMO/diversity antenna," *Prog. Electromag. Res. C*., vol. 80, pp.21–37, 2018, doi: 10.2528/PIERC17090702.

JOURNAL OF SOUTHERN COMMUNICATION ENGINEERING
ISLAMIC AZAD UNIVERSITY BUSHEHR BRANCH

E-ISSN: 2980-9231
https://sanad.iau.ir/journal/jce

https://doi.org/10.30495/jce.2023.1990276.1222

Vol. 14/ No. 56/2025

Research Article

Improving the Performance of Heart Disease Diagnosis by Combining the Cochleogram Transformation and Variable Auto-Encoder (VAE) Network

Mahbubeh Bahreini [1] | Ramin Barati[2] | Abbas Kamali[3]

[1]Department of Electrical Engineering, Shiraz Branch, Islamic Azad University, Shiraz, Iran, mahbubehbahryni@yahoo.com

[2]Department of Electrical Engineering, Shiraz Branch, Islamic Azad University, Shiraz, Iran, barati.ramin@aut.ac.ir

[3]Department of Electrical Engineering, Shiraz Branch, Islamic Azad University, Shiraz, Iran, abbas_kamali@yahoo.com

Correspondence
Ramin Barati, Assistant Professor of Electrical Engineering, Shiraz Branch, Islamic Azad University, Shiraz, Iran, barati.ramin@aut.ac.ir

Main Subjects:
Heart Disease Diagnosis with deep learning

Paper History:
Received: 12 September 2023
Revised: 13 November 2023
Accepted: 12 December 2023

Abstract

Early detection of abnormal heart sounds can largely prevent sudden death caused by heart diseases. A low-cost and non-invasive method for detecting abnormal heart sounds is the use of PCG signals. In this article, after segmenting the heart sound signal, a two-dimensional representation of the signal is obtained through cochleogram transformation. Then, with the help of deep learning and a variational autoencoder network, four final features are extracted from each signal. Finally, support vector machine and k-nearest neighbor algorithms with k-fold validation are used to classify the heart sound signal into one of the predetermined categories: normal and abnormal sound classes. In this research, the Physionet database, which contains 3,482 heart sounds from a standard collection, is used to train and evaluate the proposed method. The best results of the proposed method for classifying the two classes of heart sounds are 99.55%, 98.75%, and 99.70% in terms of accuracy, sensitivity, and specificity, demonstrating the superior capability of the proposed method compared to other methods. This abnormal sound detection system can be highly useful in rural health centers and small hospitals to assist doctors without specialized expertise in diagnosing heart problems.

Keywords: Cochleogram Transformation, Heart Disease Diagnosis, Variable Auto-Encoder Network, Support Vector Machine, K-Nearest Neighbor.

Highlights

- Gammatone filter bank extracts frequency info via cochleagram transformation.
- Latent space from autoencoders provides effective, compact features for classification.
- Autoencoders excel in unsupervised learning by compressing features efficiently.
- Robust to noise, autoencoders improve accuracy using compact latent features.

Citation: M. Bahreini, R. Barati, and A. Kamali, "Improving the Performance of Heart Disease Diagnosis by Combining the Cochleogram Transformation and Variable Auto-Encoder (VAE) Network," *Journal of Southern Communication Engineering*, vol. 14, no. 56, pp. 42-57, 2025, doi:10.30495/jce.2023.1990276.1222 [in Persian].

1. Introduction

Cardiovascular disease remains one of the most significant and concerning fatal conditions worldwide. Published statistics indicate that 32% of global deaths in 2019 were attributed to cardiovascular diseases [1]. The rising prevalence of cardiovascular disorders, their complications, and the associated healthcare costs have spurred extensive research into early detection and diagnostic methods. While various diagnostic approaches have been proposed, most are either prohibitively expensive or invasive in nature. Invasive procedures not only cause patient discomfort but may also pose life-threatening risks.

Among non-invasive, cost-effective diagnostic methods, cardiac auscultation using a stethoscope stands out for early heart disease detection. However, this method's accuracy heavily depends on the clinician's expertise [2]. Consequently, the stethoscope remains the most frequently used medical device, while phonocardiography (PCG) - the graphical recording of heart sounds - has emerged as an effective diagnostic tool for cardiovascular conditions. Healthy heart sounds typically consist of two distinct components (S1 and S2), whereas abnormal sounds result from turbulent blood flow through cardiac valves, caused by conditions such as valvular regurgitation, ventricular septal defects, or stenotic valves. While most abnormal sounds suggest cardiac pathology, some may not indicate specific diseases. These sounds can be categorized based on their spatial characteristics.

Despite its utility, PCG has limitations in detecting certain mechanical cardiac dysfunctions, including vascular anomalies or septal defects, and cannot identify coronary artery occlusions. While electrocardiography (ECG) serves as the gold standard for diagnosing mechanical heart failure, it remains relatively costly [2,3] and requires trained specialists for image acquisition, expert interpretation, and precise analysis. A highly accurate, affordable automated system could provide a reliable alternative for clinicians in resource-limited settings, potentially reducing the need for expensive and invasive diagnostic procedures.

2. Innovation and contributions

This research presents a novel approach for early detection of abnormal heart sounds using Phonocardiogram (PCG) signals. The key innovations include:

- Cochleogram Transformation: The use of cochleogram transformation to generate two-dimensional representations of heart sound signals, providing a new analytical approach.
- Deep Feature Extraction: A combination of deep learning and variational autoencoder network to extract four discriminative features from each signal, enhancing detection accuracy.
- Advanced Classification: Implementation of support vector machines and k-nearest neighbor algorithms with k-fold validation, achieving exceptional performance (accuracy: 99.55%, sensitivity: 98.75%, specificity: 99.70%).
- Comprehensive Validation: Evaluation using the standard Physionet database containing 3,482 heart sounds, demonstrating reliability and effectiveness.
- Clinical Application: A low-cost, non-invasive system suitable for rural healthcare settings, enabling general practitioners to diagnose cardiac abnormalities without specialist expertise.

These advances establish new standards in heart sound analysis, outperforming existing methods and improving cardiac healthcare outcomes.

3. Materials and Methods

Figure I illustrates the complete workflow of the proposed algorithm. The method involves several key steps: initial preprocessing followed by application of the cochleogram transform to each signal window using a gammatone filter bank. The resulting vector from the cochleogram transform is then fed into the network, where its hidden space representation serves as effective features. These extracted features are subsequently input to the classification function for final class determination. Further details of the proposed methodology are elaborated in the following sections.

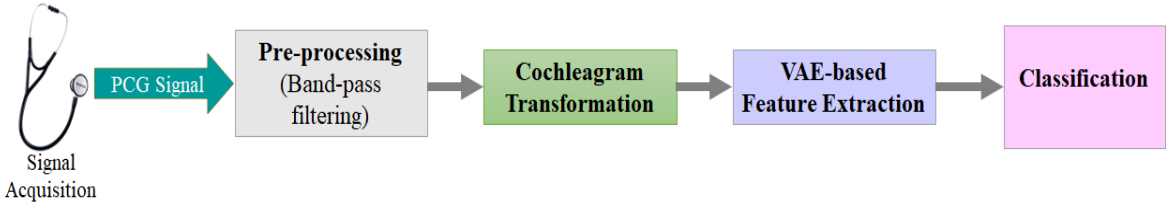


Figure I. Steps of the Proposed Method

This study utilizes the PhysioNet-2016 database for validating the proposed heart sound recognition algorithm [4,5]. The database comprises 2,435 recordings of both normal and abnormal heart sounds. While collected using

various recording devices, all signals were uniformly resampled at 2 kHz following bandwidth limitation through an anti-aliasing filter based on Nyquist-Shannon sampling theory.
The abnormal recordings primarily represent cases of valvular disorders and coronary artery occlusions, though detailed clinical annotations are limited. From these recordings, we extracted 3,482 heart sound cycles, consisting of:

- 1,754 normal cycles
- 1,728 abnormal cycles

The dataset was partitioned with 80% allocated for classifier training (organized into 128 clusters) and 20% reserved for testing. During training, 30% of the training data served as validation, with this subset being randomly selected and refreshed each epoch to ensure robust model evaluation.

4. Results and Discussion

The results demonstrate the superior performance of the proposed variational autoencoder method for dimensionality reduction when compared to other advanced feature selection techniques including Genetic Algorithm (GA), Evolutionary Algorithm (EA), and Non-Dominated Sorting Multi-Objective Symbiotic Organism Search (NSMOSOS). The key findings reveal:

- Enhanced Accuracy: The proposed variational autoencoder achieves consistently higher accuracy with both KNN and SVM classifiers, surpassing the performance of classical autoencoder and other feature selection methods.
- Improved Robustness: Combined with SVM classification, the method demonstrates significant performance gains, exceeding 99% across all key metrics (specificity, sensitivity, and accuracy).
- Effective KNN Performance: While SVM delivers superior overall results, the variational autoencoder still substantially enhances KNN classification compared to baseline and alternative approaches.

These results collectively validate the effectiveness and robustness of the proposed method for both dimensionality reduction and classification tasks, establishing its advantage over existing techniques.

5. Conclusion

This paper presents a novel approach for heart sound classification using a variational autoencoder network to extract features from cochleogram-transformed signals. Our methodology employs a balanced dataset containing 3,482 normal and abnormal PCG cycles for training and evaluation. The system processes signals through a gammatone filter bank that generates 64-dimensional vectors per frame, accurately modeling human auditory frequency response. This proven technique provides precise cochleogram feature extraction for cardiac sound analysis.
The core innovation involves our designed variational autoencoder network, which effectively reduces dimensionality by extracting 10 discriminative features from its hidden layer representation. Experimental results demonstrate exceptional classification performance, achieving 99.55% accuracy, 98.75% sensitivity, and 99.70% specificity when paired with an SVM classifier.
This system offers significant clinical potential as both a diagnostic aid for physicians and a foundation for portable heart disease detection systems. The model's computational efficiency enables deployment on standard personal computers, suggesting practical implementation in clinical settings through relatively simple hardware requirements. These capabilities position our approach as a valuable tool for non-invasive cardiac abnormality detection.

References

[1] Cardiovascular diseases (CVDs): WHO Fact Sheet. WHO (11 June 2021). https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).

[2] E. Etchells, C. Bell, K. Robb, “Does this patient have an abnormal systolic murmur?,” *Jama*, vol. 277, no 7, pp. 564–571, 1997.

[3] S. Das, S. Pal, and M. Mitra, “Deep learning approach of murmur detection using Cochleagram,” *Biomedical Signal Processing and Control*, vol. 1, no. 77, p. 103747, 2022, doi: 10.1016/j.bspc.2022.103747.

[4] C. Liu *et al.*, “An open access database for the evaluation of heart sound algorithms,” *Physiol. Meas.*, vol. 37, no. 12, pp. 2181–2213, 2016, doi: 10.1088/0967-3334/37/12/2181.

[5] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J. E. Mietus, G.B. Moody, C.-K. Peng, and H.E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, 2000, pp. 215-220, doi: 10.1161/01.cir.101.23.e215.

Declaration of Competing Interest: Authors do not have conflict of interest. The content of the paper is approved by the authors.

Author Contributions:
Mahbubeh Bahreini: Software, methodology, writing original draft preparation; **Ramin Barati:** Resources; **Abbas Kamali:** manuscript editing.

E-ISSN: 2980-9231
https://sanad.iau.ir/journal/jce

مهندسی مخابرات جنوب
دانشگاه آزاد اسلامی واحد بوشهر

https://doi.org/10.30495/jce.2023.1990276.1222

مقاله پژوهشی

بهبود عملکرد تشخیص بیماری قلبی با ترکیب تبدیل کوچلیگرام و شبکه خود رمزنگار

محبوبه بحرینی[۱] | رامین براتی*[۲] | عباس کمالی[۳]

[۱]گروه مهندسی برق، واحد شیراز، دانشگاه آزاد اسلامی، ایران.
mahbubehbahryni@yahoo.com

[۲]گروه مهندسی برق، واحد شیراز، دانشگاه آزاد اسلامی، ایران.
barati.ramin@aut.ac.ir

[۳]گروه مهندسی برق، واحد شیراز، دانشگاه آزاد اسلامی، ایران.
Ab.Kamali@iau.ac.ir

نویسنده مسئول:
*رامین براتی، استادیار گروه مهندسی برق، واحد شیراز، دانشگاه آزاد اسلامی، ایران.
barati.ramin@aut.ac.ir

موضوع اصلی:
تشخیص بیماری قلبی با الگوریتم یادگیری عمیق

تاریخچه مقاله:
تاریخ دریافت: ۲۱ شهریور ۱۴۰۲
تاریخ بازنگری: ۲۲ آبان ۱۴۰۲
تاریخ پذیرش: ۲۱ آذر ۱۴۰۲

چکیده:

تشخیص اولیه صداهای غیرطبیعی قلب تا حد زیادی می‌تواند از مرگ ناگهانی ناشی از بیماری‌های قلبی جلوگیری کند. یک روش کم‌هزینه و غیر هجومی برای تشخیص صداهای غیرطبیعی قلب، به‌کارگیری سیگنال PCG است. در این مقاله، بعد از قسمت‌بندی سیگنال صدای قلب، نمایش دوبعدی سیگنال توسط تبدیل کوچلیگرام حاصل می‌شود، سپس به کمک یادگیری عمیق و شبکه خود رمزنگار متغیر، چهار ویژگی نهایی از هر سیگنال استخراج می‌شود. در نهایت از ماشین بردار پشتیبان و k-نزدیک‌ترین همسایه با اعتبارسنجی K-fold برای طبقه‌بندی سیگنال صدای قلب، در یکی از دسته‌های از پیش تعیین‌شده کلاس صدای نرمال و غیرنرمال استفاده می‌شود. در این پژوهش از مجموعه داده‌ی فیزیونت که دارای ۳۴۸۲ صدای قلب از یک مجموعه استاندارد است، جهت آموزش و ارزیابی روش پیشنهادی استفاده می‌شود. بهترین نتایج به‌دست‌آمده روش پیشنهادی جهت طبقه‌بندی دو کلاسه صدای قلب به‌ترتیب به‌دقت، حساسیت و ویژه بودن۹۹/۵۵، ۹۸/۷۵ و ۹۹/۷۰ است که توانایی بالاتر روش پیشنهادی در مقایسه با سایر روش‌های موجود را اثبات می‌کند. از این سیستم تشخیص صداهای غیرنرمال می‌توان به‌صورت بسیار مفید در مراکز درمانی بهداشت روستایی و بیمارستان‌های کوچک به‌منظور کمک به پزشکان بدون تخصص برای تشخیص مشکلات قلبی استفاده کرد.

کلید واژه‌ها: تبدیل کوچلیگرام، تشخیص بیماری قلبی، شبکه خود رمزنگار متغیر، ماشین بردار پشتیبان، K-نزدیک‌ترین همسایه

تازه های تحقیق:

- بانک فیلتر گاماتون اطلاعات فرکانس را از طریق تبدیل حلزون گوش استخراج می‌کند.
- فضای پنهان از رمزگذارهای خودکار ویژگی های موثر و فشرده را برای طبقه بندی فراهم می‌کند.
- رمزگذارهای خودکار در یادگیری بدون نظارت با فشرده سازی کارآمد ویژگی ها برتری می‌یابند.
- رمزگذارهای خودکار که دربرابر نویز مقاوم هستند، دقت را با استفاده از ویژگی‌های پنهان فشرده بهبود می‌بخشند.

۱-مقدمه

بیماری قلبی عروقی یکی از مهم‌ترین بیماری‌های مهلک و نگران‌کننده در کل دنیا است. بر اساس آمارهای منتشر شده ۳۲ درصد از مرگ و میر در کل دنیا در سال ۲۰۱۹ ناشی از بیماری‌های قلبی عروقی بوده است [۱]. رشد روزافزون بیماری‌های قلبی و عروقی و عوارض آن‌ها و هزینه‌هایی که به سیستم پزشکی جامعه تحمیل می‌کند باعث شده است که تحقیقات زیادی در زمینه بررسی روش‌های شناسایی و تشخیص زودهنگام آن صورت پذیرد. اگرچه روش‌های تشخیص و شناسایی متفاوتی ارائه شده است، اما اکثر آن‌ها یا بسیار گران است و یا ماهیت هجومی دارد. روش‌های هجومی هم می‌توانند باعث ناراحتی بیمار شده و هم برای او مرگ‌آفرین باشند. یکی از روش‌های غیر هجومی، ارزان و ساده برای تشخیص اولیه بیماری قلبی، شنیدن صدای قلب با استفاده از گوشی پزشکی است. هرچند که دقت این رویکرد بستگی زیادی به مهارت پزشک دارد [۲]. از این رو استتوسکوپ بیش از هر وسیله پزشکی مورد استفاده پزشکان قرار می‌گیرد و فونوکاردیوگرافی(PCG)[1]و یا ثبت گرافیکی سیگنال صدای قلب به‌عنوان یک راهکار مؤثر به منظور تشخیص بیماری‌های قلبی عروقی شناخته شده است. صدای قلب سالم دارای دو صدای واضح و مجزا به نام‌های S1 و S2 می‌باشد و صداهای غیرعادی قلب، صدای گذر ناگهانی جریان خون از درون دریچه قلب است که ناشی از بازگشت خون از طریق دریچه سوراخ‌دار، نقص جدار بطنی و یا جریان رو به جلو از طریق دریچه‌های باریک است. اکثرا این صداهای مبهم غیرطبیعی نشان‌دهنده مشکلات قلبی بوده، اما بعضاً نشان از هیچ مورد خاص بیماری نیست. صداهای غیرطبیعی قلب را می‌شود بر اساس مشخصه‌های فضایی آنها دسته‌بندی کرد. اگرچه PCG یک روش ثبت گرافیکی سیگنال یک روش تشخیص بیماری قلبی است، اما عدم کارکرد مکانیکی قلب، شامل ایرادات عروقی یا نواقص SEPTAL ، با این روش قابل‌تشخیص نمی‌باشد. به طوری که دستگاه PCG قابلیت تشخیص انسداد عروق کرونری را ندارد. اگرچه ECG [2] یک روش استاندارد برای تشخیص نارسایی مکانیکی قلب است، اما به‌نوعی یک روش پرهزینه است [۲، ۳]. همچنین نیاز به یک کارشناس حرفه‌ای و آموزش‌دیده برای اخذ تصاویر، ارائه نظر کارشناسی و تفسیر صحیح و دقیق دارد. سیستم خودکار بادقت بالا و هزینه کم، می‌تواند جایگزین قابل‌اعتمادی برای پزشکان در بیمارستان‌های کوچک باشد. همچنین این روش نیاز به روش‌های تشخیصی پرهزینه و تهاجمی را به‌شدت کاهش می‌دهد.

در طول یک دهه اخیر، به دلیل پیشرفت‌های حاصل در زمینه هوش مصنوعی، محققان به آنالیز و پردازش سیگنال PCG توسط روش‌های یادگیری ماشین به‌شدت علاقه‌مند شده‌اند و تعداد قابل‌توجهی از تحقیقات بر روی شناسایی، تشخیص و دسته‌بندی صداهای طبیعی و غیرطبیعی قلب انجام شده است، این تحقیقات بر اساس ویژگی‌های استخراج شده از حوزه‌ی فرکانس [۳۲،۶،۴] و حوزه‌ی زمان-فرکانس است [۷-۱۵،۱۴،۱۲]. دلیل اصلی انتخاب ویژگی در حوزه‌ی زمان-فرکانس، رفتار غیرایستا و شبه‌تناوبی سیگنال‌های PCG است.

سافارا و همکاران، با استفاده از ترکیب تبدیل موجک بسته‌ای[3] و استخراج ویژگی مبتنی بر آنتروپی[4]، راهکاری را برای دسته‌بندی صدای قلب بادقت ۹۶/۹۴ درصد پیشنهاد دادند [۷]. برای تشخیص صدای غیر طبیعی قلب (سیستولیک) روش ترکیب آنتروپی ضرایب موجک بسته‌ای با استفاده از توان ضرایب موجک [۸] پیشنهاد شده است. علت رایج استفاده از تبدیل موجک و انواع آن در تحلیل و آنالیز سیگنال PCG، امکان دسترسی به محتوای زمان-فرکانس و تجزیه سیگنال بر اساس فرکانس مشخص (فرکانس پایین و فرکانس بالا) است. روش تبدیل موجک پیوسته[5] (CWT)، بر اساس موجک مورلت[6]، نمونه‌ی دیگری است که برای تشخیص صداهای غیرطبیعی قلب پیشنهاد شده است [۹،۱۰]. وارگیز و همکاران، روش تبدیل موجک تجربی[7] (EWT) را برای آنالیز سیگنال صدای قلب پیشنهاد دادند [۱۱]. EWT روشی است که ابتدا اجزاء فرکانس را پیدا می‌کند و سپس روش فیلترینگ را بر روی آن پیاده می‌کند. این فرایند، ترکیبی از تجزیه مد تجربی (EMD) و روش تجزیه تطبیقی تبدیل موجک (WT) است.

[1] Phonocardiography
[2] Electrocardiogram
[3] Packet wavelet transform
[4] Entropy
[5] Continuous wavelet transform
[6] Morelet wave
[7] Experimental wavelet transform

کیفیت تأثیر روش‌های مبتنی برWT وابسته به دو فاکتور مهم موجک مادر و تعداد سطوح تجزیه است. از دیگر ابزار پرکاربرد ویژگی حوزه‌ی زمان-فرکانس، ویژگی‌های ضرایب کپسترال فرکانسی مل[۱] (MFCC) است [۱۲]. بعد از موفقیت چشم‌گیر استفاده از این روش در تجزیه‌وتحلیل گفتار و صدا، از آن برای آنالیز و بررسی سیگنال صدای قلب نیز استفاده شده است. ویژگی صوتی برگرفته از ضرایب کپسترال مل فرکانسی برگرفته از سیستم غیرخطی ادراک شنوایی انسان است و دارای ۱۳ بعد است [۱۳]. مکینتاش و همکاران، از روش MFCC برای دسته‌بندی سیگنال صدای طبیعی و غیرطبیعی قلب استفاده کرده‌اند [۱۴و۳۰]. علاوه بر MFCC برای استخراج مجموعه‌ای از ویژگی‌های فراکتال از روش‌هایی چون برازش منحنی استفاده شده است [۱۵]. ویژگی MFCC، در حضور اعوجاج کم عملکرد خوبی دارد.

افت فاحش عملکرد این روش زمانی است که صداهای ضبط شده قلب تا حدودی با اغتشاشات و نویز آمیخته باشد که البته باتوجه‌به روش‌های داده‌برداری بدون کنترل قابل‌درک است. روش‌های یادگیری ماشین به‌منظور تجزیه‌وتحلیل PCG به ترتیب شامل مراحل استخراج ویژگی‌ها، انتخاب ویژگی‌ها و دسته‌بندی ویژگی‌ها هستند. محققان بر اساس ویژگی‌های انتخابی‌شان از انواع مختلف دسته‌بندی‌کننده‌ها استفاده می‌کنند. در سال‌های اخیر اکثر محققین ترجیح می‌دهند از روش‌های یادگیری عمیق برای دسته‌بندی صدای قلب استفاده کنند. در حال حاضر شبکه عصبی عمیق یکی از روش‌های پراستفاده و مورد استقبال محققان در زمینه‌های مختلف است. در تحقیقات زیادی، از شبکه عصبی عمیق (DNN) برای دسته‌بندی صدای قلب به دسته نرمال و غیرنرمال استفاده شده است [۱۲، ۱۷-۱۵]. کریشنان و همکاران، برای دسته‌بندی صداهای قلب از ۱۰۸۱ نمونه ثبت شده PCG در مجموعه دیتا عمومی، برای آموزش و اعتبارسنجی مدل‌های DNN پیشنهادی استفاده کرده است [۱۸]. مدل WAVENET عمیق برای دسته‌بندی ۵ نوع از صداهای غیرطبیعی قلب پیشنهاد شده است [۱۹]. DNN فقط قادر به دسته‌بندی دیتاهای سری زمانی است درحالی‌که CNN می‌تواند دیتاهای سری زمانی را به تصاویر تبدیل کند و به‌عنوان لایه ورودی CNN به کار گیرد. [۲۱،۲۰،۱۴] مدل‌های CNN به ترتیب در MFCC، تبدیل موج متقابل[۲] و سنسورهای شنوایی نورومورفیک[۳] (NAS) استفاده شده است. این ویژگی‌ها برای توصیف سیگنال صدای قلب به‌صورت تصویر مورد استفاده قرار گرفته است. همچنین در [۱۲] قابلیت بالای تبدیل کوچلیگرام[۴] برای طبقه‌بندی سیگنال PCG و بدون اعمال هیچ‌گونه فیلتری (پیش‌پردازش) نشان‌داده شده است [۳۱،۲۹]. بانک فیلتر گاماتون[۵] با الگوبرداری از مدل بانک فیلتر سیستم شنوایی انسان، ویژگی‌های کوچلیگرام را استخراج کرده است. با استفاده از ماتریس ویژگی ۶۴ بعدی کوچلیگرام، یک مدل DNN، برای دسته‌بندی صداهای طبیعی و غیرطبیعی قلب آموزش‌داده شده است [۳].

در ادامه بخش‌بندی مقاله به‌صورت زیر است. مواد و روش‌های مورداستفاده در این کار در بخش۲ توضیح داده شده است. مجموعه‌داده‌ها، فرایند استخراج ویژگی و مدل دسته‌بندی در بخش۲ به طور خلاصه ارائه شده است. آزمایش‌ها و نتایج دقیق در بخش ۳ گزارش شده است. در بخش ۴ خلاصه مباحث ذکرشده است و در نهایت، بخش ۵ به نتیجه‌گیری تحقیق می‌پردازد.

۲- مواد و روش‌ها

۲-۱- تبدیل کوچلیگرام

به‌طورکلی تبدیل‌های زمان-فرکانس را می‌توان به دودسته پویا و ایستا گروه‌بندی کرد. در روش‌های پویا تلاش می‌شود که زمان به‌صورت واقعی و نه انتزاعی نمایش داده شود. تبدیل کوچلیگرام در واقع همان اسپکتروگرام پویا یا اسپکتروگرام با مدل حلزونی گوش است و در آن از ضرایب کپسترال فرکانس گاماتون[۶] که یک ساختار فیلتر بانکی دارد برای استخراج اطلاعات فرکانسی استفاده می‌شود. ضرایب کپسترال فرکانس گاماتون مشابه ضرایب کپسترال فرکانس مل[۷] است و تنها تفاوت آن در نوع فیلتر و پهنای باند مربوط به آن است، به طوری که فیلتر استفاده شده در فیلتر گاماتون به کمک فیلترهای حلزونی مدل شده است.

[1] Mel frequency capstral coefficients
[2] Coss-Wavelet Transform
[3] Nerumorphic Auditory Sensors
[4] Kochilgram
[5] Gammatone filter bank
[6] Gammatone Frequency Capstral
[7] Mel frequency

برای محاسبه ضرایب کپسترال فرکانسی گاماتون، اندازه‌ی پاسخ فرکانسی به‌دست‌آمده از اعمال تبدیل فوریه بر روی سیگنال صدای قلب، از فیلتر بانک گاماتون که به صورت رابطه ۱ تعریف می‌شود عبور داده می‌شود،

$$G(\omega)=\frac{A}{2}\frac{6}{(2\pi b\,\mathrm{ERB}(f_c))+j(\omega-\omega_c)^4}+\frac{A}{2}\frac{6}{(2\pi b\,\mathrm{ERB}(f_c))-j(\omega-\omega_c)^4} \quad (۱)$$

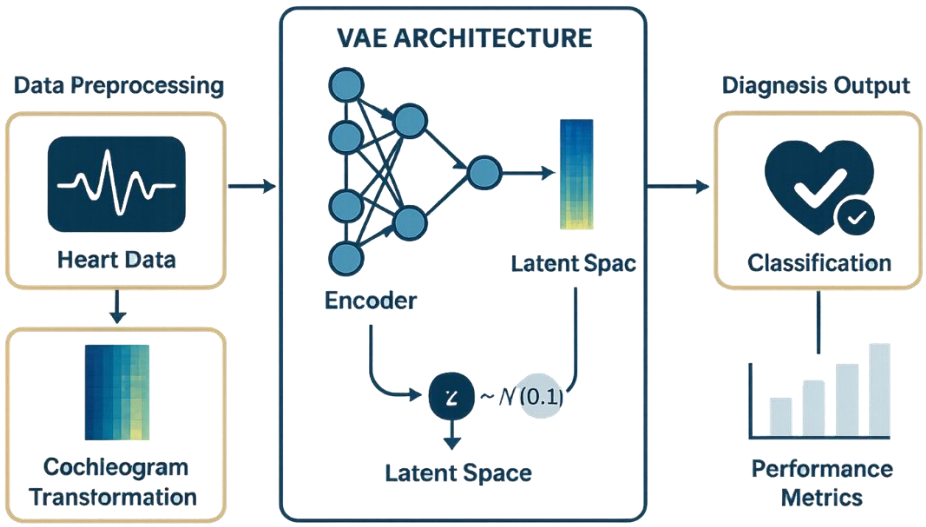
که در رابطه بالا ERB پهنای باند مستطیلی متناظر با فرکانس است. که به کمک آن می‌توان طیف هموارتری برای دامنه انتظار داشت. در نهایت لگاریتم طیف سیگنال‌های میان‌گذر به حوزه‌ی زمان انتقال داده می‌شود. ضرایب بدست آمده، ضرایب کپستروم گاماتون خواهد بود. بیان طیفی سیگنال صدای قلب به صورت ضرایب کپستروم گاماتون از نظر ویژگی‌های طیفی محلی می‌تواند قابلیت بالایی در تمایز کلاس‌های مختلف داشته باشد. با توجه به اینکه ضرایب طیف گاماتون در نتیجه لگاریتم اعداد حقیقی هستند به کمک تبدیل کسینوسی گسسته می‌توان سیگنال حوزه‌ی زمان را بازیابی کرد. از این‌رو اگر ضرایب طیفی توان با S_k به ازای k=1,2, …, K نشان داده شود، ضرایب کپسترال فرکانس گاماتون به صورت رابطه ۲ بدست می‌آید،

$$c_n=\sum_{k=1}^{K}(\log S_k)\cos\left(n\left(k-\frac{1}{2}\right)\frac{\pi}{k}\right),\ n=1,2,...,k \quad (۲)$$

در این مقاله از ۶۴ فیلتر گاماتون با ضریب پهنای باند ۱۵ استفاده شده است.

۲-۲- ساختار روش پیشنهادی

در شکل ۱ تمام مراحل الگوریتم پیشنهادی به طور خلاصه نشان‌داده‌شده است. در روش پیشنهادی، بعد از پیش‌پردازش اولیه، تبدیل کوچلیگرام سیگنال با به‌کارگیری بانک فیلتر گاماتون بر روی هر پنجره از سیگنال اعمال می‌شود و سپس بردار به‌دست‌آمده از تبدیل کوچلیگرام وارد شبکه خود رمزنگار می‌شود و از فضای پنهان آن به‌عنوان ویژگی مؤثر استفاده می‌شود و در نهایت ویژگی‌های به‌دست‌آمده وارد تابع دسته‌بند می‌شوند و کلاس داده مشخص می‌گردد. در ادامه جزئیات بیشتری در مورد روش پیشنهادی ارائه می‌شود.



شکل ۱: مراحل روش پیشنهادی
Figure 1. Steps of the Proposed Method

۲-۲-۱- پایگاه داده

در این بخش از پایگاه‌داده PhysioNet-2016 برای اعتبارسنجی الگوریتم پیشنهادی مبتنی بر تشخیص صدای قلب استفاده شده است [۲۴،۲۳] . این پایگاه‌داده شامل ثبت ۲۴۳۵ سیگنال صدای قلب شامل صداهای طبیعی و غیرطبیعی قلب است. اگرچه فرایند دستیابی به داده‌ها با استفاده از دستگاه‌های ضبط مختلف انجام شده است، اما نتایج تمام ثبت‌ها در فرکانس ۲KHz با استفاده از فیلتر تصحیح فرکانس دوباره نمونه‌برداری می‌شوند (فیلتر تصحیح فرکانس فیلتری است که قبل از نمونه

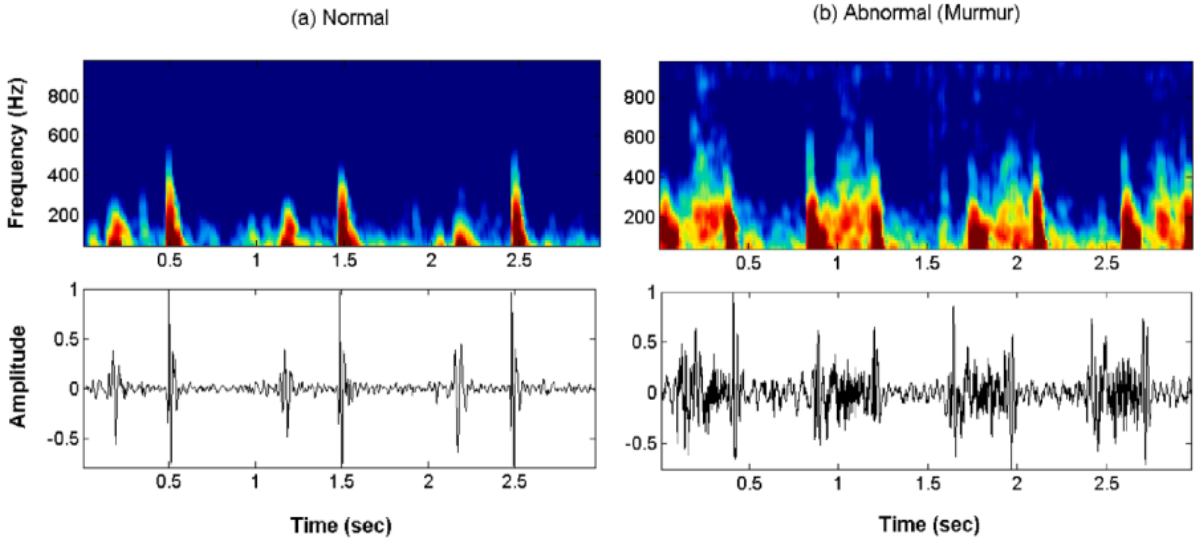
برداری سیگنال برای محدود کردن پهنای باند یک سیگنال مورد استفاده قرار می‌گیرد و از تئوری نمونه برداری نایکوئیست شانون استفاده می‌کند). بیشتر صداهای غیر طبیعی قلب از بیماران مبتلا به اختلالات دریچه قلب و انسداد عروق کرونری جمع‌آوری می‌شوند و جزئیات دیگری در دسترس نیست. با توجه به سیگنال‌های موجود یک مجموعه داده شامل ۳۴۸۲ سیکل از صدای قلب بدست آمده که شامل دو کلاس صدای نرمال (۱۷۵۴) و کلاس غیرنرمال (۱۷۲۸) می‌باشد و از ٪۸۰ از کل داده‌ها برای آموزش و ٪۲۰ برای تست دسته‌بندی کننده استفاده شده است . داده‌های آموزشی به دسته‌های ۱۲۸تایی تقسیم شده‌اند. ٪۳۰ از داده‌های آموزشی برای اعتبارسنجی نگه‌داشته شده که به صورت تصادفی انتخاب و در هر دوره تغییر داده شده است.

۲-۲-۲-پیش‌پردازش داده

پایگاه داده صداهای قلب PhysioNet از سراسر جهان با استفاده از سیستم‌های مختلف دستیابی داده در محیط‌های بالینی و غیربالینی جمع‌آوری می‌شوند، بنابراین در ابتدا دامنه داده‌های PCG با استفاده از تکنیک Min-Max نرمالیزه شده است.

۲-۲-۳- استخراج ویژگی اولیه با اعمال تبدیل کوچلیگرام

بعد از پیش‌پردازش اولیه، فرایند استخراج ویژگی انجام می‌شود. در گام اولیه فرایند استخراج ویژگی، داده‌های PCG با طول فریم ۶۴ میلی‌ثانیه و با جابجایی زمانی ۱۰ میلی‌ثانیه برای محاسبه تبدیل فوریه زمان کوتاه استفاده شده است. به‌عنوان‌مثال، اگر فاصله زمانی صدای قلب ۱۵۸ میلی‌ثانیه در نظر گرفته شود، درکل شامل ده فریم خواهد بود. نمونه‌های با طول ۴ میلی‌ثانیه واقع در بخش انتهایی حذف خواهند شد. هر داده اضافی فراتر از چارچوب نهایی که کمتر از ۱۰ میلی ثانیه است، کنار گذاشته می‌شود. سپس از بانک فیلتر گاماتون ۶۴ فیلتری، دارای فرکانس‌های مرکزی در محدوده ۵۰ هرتز تا ۱ کیلوهرتز برای به دست آوردن ماتریس ویژگی ۶۴ بعدی مربوط به تبدیل کوچلیگرام استفاده می‌شود. در شکل ۲ به عنوان نمونه یک سیگنال صوتی صدای نرمال و غیرنرمال قلب و کوچلیگرام مربوط به آن نشان داده شده است.

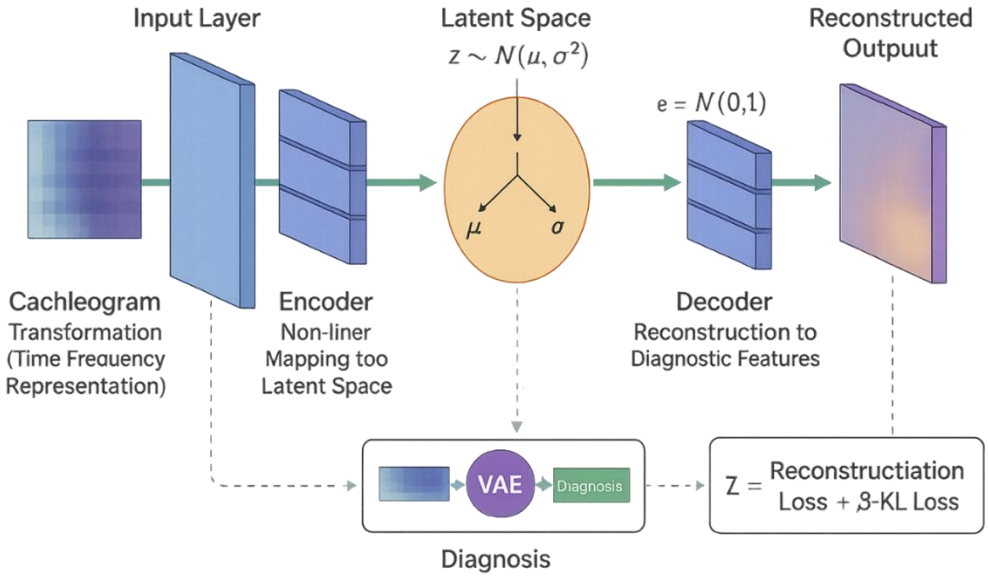


شکل ۲: نمونه سیگنال‌های صوتی طبیعی و غیر طبیعی قلب و خروجی مربوط به تبدیل کوچلیگرام آن
Figure 2. Samples of normal and abnormal sound signals of the heart and the output related to its cochleogram conversion

۲-۲-۴- استخراج ویژگی با خودرمزنگار عمیق

خودرمزنگار نقش مهمی در یادگیری بدون نظارت ایفا می کند. بازسازی داده‌ها هدف اصلی این نوع شبکه‌ها می‌باشد و اساسی‌ترین مزیت این شبکه‌ها استخراج ویژگی‌های فشرده می‌باشد. به بیان دیگر خودرمزنگارها کارایی بسیار مناسبی در کاهش ابعاد و استخراج ویژگی و بازنمایی داده‌ها دارند و اخیراً مورد توجه پژوهشگران زیادی قرار گرفته است. بعد از اعمال

تبدیل کوچلیگرام،ابعاد داده‌ی مسئله ۶۴ می‌باشد که افزونگی بسیاری دارد و برای حل مسئله کلاس‌بندی به تمامی مقادیر داده‌ها احتیاج نداریم با توجه به اینکه تعداد داده‌های آموزشی محدود می‌باشد و برای بهبود عملکرد سیستم طبقه‌بندی و همچنین کاهش پیچیدگی محاسباتی نیاز به کاهش ابعاد داده‌ی ورودی داریم ، ساختار خودرمزنگار این امکان را فراهم می‌کند که با طراحی یک شبکه عمیق، بهترین ویژگی‌ها از طریق لایه پنهان (latent) از هر سیگنال استخراج شود.



شکل ۳: ساختار شبکه خودرمزنگار
Figure 3. Self-encrypting network structure

همانطور که در شکل ۳ مشخص است، ساختار کلی خودرمزنگارها به دو بخش رمزگذار و رمزگشا تقسیم می‌شود. در قسمت رمزگذار داده‌های ورودی (خروجی فیلتربانک گاماتون) را به فضای پنهان $\mathbf{z} \in \mathbb{R}^L; L < F$ نگاشت می کند (F بعد فضای ورودی و L بعد فضای پنهان است) که در این مقاله بعد فضای ویژگی ۴ در نظر گرفته شده است و در بخش رمزگشایی از فضای ویژگی ، داده‌های اصلی بازسازی می‌شود. خودرمزنگار متغیر یا VAE یک نوع جدید از خودرمزنگارها می‌باشد که برای کاهش ابعاد داده‌های نویزی از طریق فضای پنهان استفاده می شود. عملکرد کلی این شبکه مشابه یک خودرمزگذار معمولی است که مبتنی بر یادگیری بدون نظارت، یک بردار خروجی تولید کند که حدوداً مشابه بردار ورودی می‌باشد. تفاوت اصلی آن این است که فضای پنهان با پارامترهای یک توزیع احتمال مشخص می‌شود، به عبارتی خودرمزنگار متغیر مدل احتمالی خودرمزنگار کلاسیک می‌باشد. برای این منظور، تابع هزینه در خودرمزنگار متغیر ترکیبی از میانگین مربعات خطا و تابع KullbackLeibler است به همین دلیل خودرمزنگار متغیر قابلیت حذف نویز و داده‌های پرت را دارد و فضای پنهان را می‌توان به عنوان بهترین و فشرده‌ترین ویژگی‌ها برای بهبود دقت طبقه‌بندی در نظر گرفت.

۲-۲-۴-۱- هایپرپارامترهای مورد استفاده در شبکه خودرمزنگار

به منظور همگرایی سریع‌تر شبکه خودرمزنگار، چندین هایپرپارامتر وجود دارد که باید آن‌ها را به صورت سعی و خطا تنظیم کرد. در ادامه مهم‌ترین آن‌ها بررسی شده است:

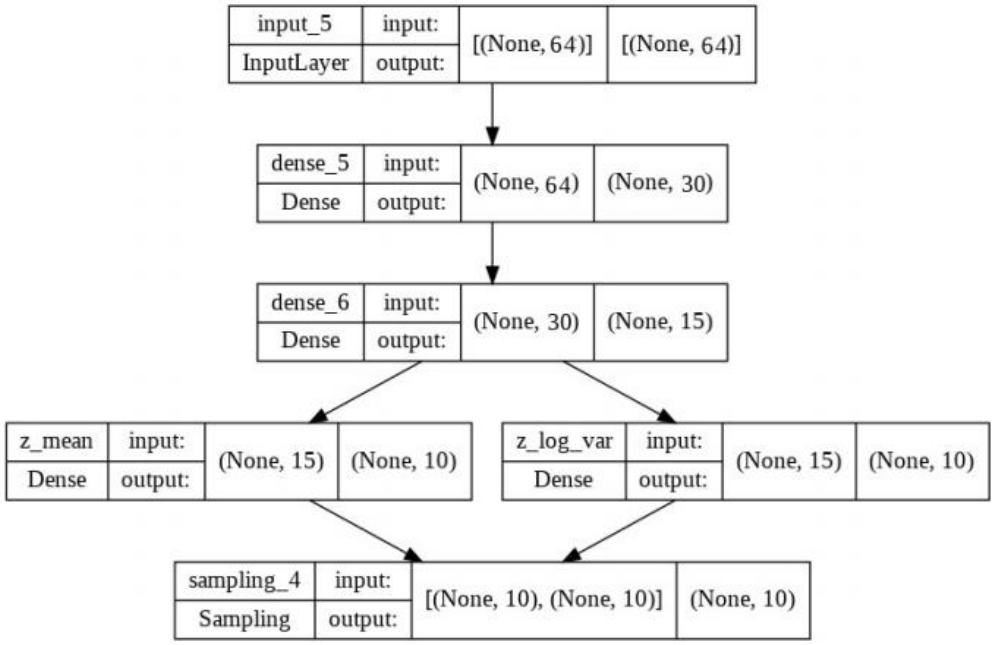
۱. ابعاد فضای پنهان: این متغیر ابعاد نهایی فضای ویژگی را مشخص می‌کند. یک فضای پنهان با ابعاد کمتر می‌تواند ویژگی‌های پیچیده‌تری از داده‌های ورودی را استخراج کند، اما ممکن است شبکه همگرا نگردد و یا دیرتر همگرا شود.

۲. نرخ یادگیری: همان گونه که از اسم این پارامتر مشخص است سرعت یادگیری توسط این پارامتر تعیین می‌گردد. نرخ یادگیری بالا سرعت یادگیری را افزایش می‌دهد، با این حال ممکن است شبکه همگرا نگردد و نرخ یادگیری کوچک، همگرایی شبکه را طولانی می‌کند.

۳. اندازه‌ی بچ[1]: این پارامتر تعداد داده‌های ورودی به شبکه را مشخص می کند. هرچقدر اندازه‌ی بچ بزرگ‌تر باشد همگرایی سریع‌تر می‌شود، اما فرایند آموزشی ثبات کمتری دارد.

۴. طراحی شبکه، تعداد لایه‌ها و نرون‌های هر لایه: این پارامتر عمق شبکه‌های رمزگذار و رمزگشا را مشخص می‌کند. هر چقدر شبکه عمیق‌تر باشد قابلیت یادگیری الگوهای پیچیده‌تر را دارد، اما ممکن است به زمان آموزشی و منابع محاسباتی بیشتری نیز نیاز داشته باشد.

با توجه به اینکه تعیین مناسب این هایپرپارامترها می‌تواند تأثیر زیادی بر همگرایی و عملکرد نهایی شبکه داشته باشد، یک‌سری آزمایش‌ها به‌صورت سعی و خطا صورت گرفته است و مقادیری که بهترین نتیجه را داشته‌اند برای الگوریتم پیشنهادی انتخاب شده است و در شکل ۴ جزئیات شبکه پیشنهادی نشان‌داده شده است. همچنین نرخ یادگیری و اندازه بچ به ترتیب ۰/۰۰۰۲ و ۱۲۸ در نظر گرفته شده است.



شکل ۴: شبکه خود رمزنگار برای استخراج ویژگی [۳۳]

Figure 4. Self-encrypting network for feature extraction

۲-۳-طبقه بندی

در نهایت، ویژگی‌های استخراج شده به کمک یک تابع دسته‌بند مشخص دسته‌بندی می‌شوند برای این‌کار توابع دسته‌بند زیادی وجود دارد که می‌توان از آن‌ها استفاده کرد، مانند تحلیل تفکیک خطی[2] (LDA)، K- نزدیک‌ترین همسایه[3] (KNN)، ماشین‌های بردار پشتیبان[4] (SVM)، جنگل‌های تصادفی و شبکه‌های عصبی[5] (NN). در این پژوهش از تابع SVM و KNN برای طبقه‌بندی نهایی استفاده می‌شود که در پژوهش‌های قبلی قابلیت این دسته‌بندها در کار با سیگنال‌های حیاتی اثبات شده است [۲۵] روشی سریع با درصد صحت بالا، طبقه‌بندی را انجام می‌دهد، بنابراین دقت عملکرد به میزان زیادی بهبود می‌یابد. در این روش با استفاده از روش‌های تکاملی، نمونه‌ها می‌توانند با حداقل مجموع فواصل درون خوشه‌ای، خوشه‌بندی شوند و به منظور کاهش خطا در طبقه‌بندی، خوشه‌ها در دورترین مکان نسبت به هم واقع شوند. همچنین به دلیل اینکه هایپرپارامترهای تابع دسته‌بند بر روی عملکرد نهایی سیستم تأثیرگذار است، یک سری بهینه‌سازی بر روی هایپرپارامترهای تابع دسته‌بند انجام می‌شود. در

[1] Batch
[2] Linear Discriminant Analysis
[3] Nearest Neighbor
[4] Support Vector Machines
[5] Neural Networks

این مقاله یک تابع دسته‌بند SVM با هسته گاوسی γ برابر با ۰/۰۱ و C برابر با ۲ و یک تابع دسته‌بند KNN با K=۵ و بافاصله ماهالانویوس[1] در نظر گفته می‌شود. در بخش بعدی نتایج به‌دست آمده را گزارش می‌کنیم.

۲-۴-معیارهای ارزیابی

برای ارزیابی عملکرد روش پیشنهادی، بعد از تخمین ماتریس درهم‌ریختگی محاسبه، دقت (.Acc)، حساسیت (.Se) و درصد ویژه بودن (.Sp) محاسبه می‌شود. نسبت نمونه‌های پیش‌بینی‌شده درست به تعداد کل نمونه‌ها به‌عنوان دقت شناخته می‌شود که اعتبار طبقه‌بندی‌کننده را اندازه‌گیری می‌کند. حساسیت به‌عنوان نسبت نمونه‌های مثبت واقعی به همه‌ی نمونه‌های که مثبت تشخیص داده شده‌اند. ویژه بودن به‌عنوان نسبت موارد منفی واقعی به همه‌ی نمونه‌های که منفی تشخیص داده شده‌اند. روابط ریاضی برای محاسبه این معیارها عبارت‌اند از،

$$Acc.(\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (۳)$$

$$Se.(\%) = \frac{TP}{TP + FN} \times 100 \quad (۴)$$

$$Sp.(\%) = \frac{TN}{TN + FP} \times 100 \quad (۵)$$

که در آن TP و TN تعداد نمونه‌هایی هستند که به ترتیب در کلاس‌های نرمال و غیرنرمال به درستی پیش‌بینی شده‌اند. FP تعداد نمونه‌هایی است که مدل به طور نادرست به عنوان نرمال پیش‌بینی می‌کند و FN تعداد نمونه‌هایی است که مدل به اشتباه آن‌ها را غیرنرمال پیش‌بینی می‌کند. همچنین برای گزارش نهایی مقادیر، اعتبارسنجی متقابل با ۵ لایه انجام می‌شود تا اطمینان حاصل گردد که هیچ بایاسی در عملکرد طبقه‌بندی وجود ندارد. در این مرحله، مجموعه داده‌ها به پنج زیر گروه افراز می‌شوند، از این ۵ زیر مجموعه هر بار ۴ زیر مجموعه برای آموزش و زیر مجموعه باقی‌مانده برای ارزیابی و تست استفاده می‌شود و این روال ۵ بار تکرار می‌شود به‌طوری‌که هر نمونه یکبار برای ارزیابی استفاده گردد. در نهایت میانگین این ۵ بار تست به‌عنوان میزان عملکرد نهایی گزارش می‌شود.

در این قسمت ابتدا آزمایشی را برای نشان‌دادن مزیت کاهش بعد توسط شبکه خود رمزنگار متغیر[2] انجام می‌دهیم. سپس، رویکرد پیشنهادی خود را با سایر روش‌های به‌روز موجود مقایسه می‌کنیم تا مزیت و قابلیت رویکرد پیشنهادی خود را نشان دهیم. در ادامه ابتدا عملکرد رویکرد کاهش بعد مبتنی بر رویکرد پیشنهادی را با سه روش انتخاب ویژگی پیشرفته مقایسه می‌کنیم، همچنین حالتی که در آن انتخاب ویژگی صورت نمی‌پذیرد را به‌عنوان روش پایه در نظر می‌گیریم. انتخاب ویژگی مبتنی بر الگوریتم ژنتیک[3] (GA) [۲۶]، انتخاب ویژگی مبتنی بر الگوریتم‌های تکاملی[4] (EA) [۲۷]، و الگوریتم جستجوی ارگانیسم‌های هم‌زیست چندهدفه مرتب‌سازی غیرمسلط[5] (NSMOSOS) [۲۸]، برای مقایسه استفاده شده است. نتایج عددی رویکرد پیشنهادی در جدول ۱ خلاصه شده است.

در این قسمت فرض شده است که تمام مراحل (پیش‌پردازش، استخراج ویژگی، طبقه‌بندی‌کننده) یکسان هستند و تنها تفاوت در مرحله کاهش بعد و انتخاب ویژگی است، سپس رویکرد پیشنهادی را بر اساس معیارهای دقت، حساسیت و ویژه بودن مقایسه می‌کنیم. اولین و بدیهی‌ترین مشاهدات این است که استفاده از همه ویژگی‌ها کمترین عملکرد را در طبقه‌بندی دارد که مشکل ابعاد بزرگ را نشان می‌دهد. همان‌طور که می‌بینیم، روش کاهش بعد مبتنی بر یادگیری عمیق پیشنهادی، قادر است که به عملکرد بالاتری نسبت به روش‌های دیگر دست یابد. در مقایسه روش پیشنهادی مبتنی بر خود رمزنگار کلاسیک با خود رمزنگار متغیر، برتری ناچیز در روش خود رمزنگار متغیر مشخص است. برای نشان‌دادن مزیت اصلی روش خود رمزنگار متغیر حالتی را در نظر می‌گیریم که سیگنال ورودی با نویز گوسی همراه باشد. نتایج به‌دست‌آمده از این سناریو در جدول ۲ نشان‌داده‌شده است.

---

[1] Mahalanobis
[2] Autoencoder
[3] Genetic algorithm
[4] Evolutionary Algorithms
[5] Non-dominated sorting multi-objective symbiotic organisms search algorithm

جدول ۱: بررسی عملکرد شبکه خود رمزنگار متغیر در کاهش ابعاد ورودی و مقایسه آن با شبکه خود رمزنگار کلاسیک و روش‌های نوین دیگر
Table 1. Investigating the performance of the variable self-encryptor network in reducing input dimensions and comparing it with the classic self-encryptor network and other modern methods.

Ref.	Feature Reduction	Category Function	Precision	Sensitivity	Being Special
-	All features	KNN	84.22	84.38	82.19
		SVM	84.56	84.62	83.72
[26]	Genetic algorithm	KNN	89.90	90.27	90.19
		SVM	91.22	93.37	92.65
[27]	Evolutionary algorithm	KNN	96.49	98.38	97.56
		SVM	97.88	98.62	98.02
[28]	Searching for symbiotic organisms	KNN	97.19	98.25	97.41
		SVM	98.22	98.81	98.15
Proposed method	Classic self-encryption	KNN	98.87	98.29	98.23
		SVM	99.02	98.37	99.14
Proposed method	Variable self-encryption	KNN	99.43	98.62	99.21
		SVM	55	98.75	99.70

جدول ۲: بررسی عملکرد شبکه خود رمزنگار متغیر در کاهش ابعاد ورودی و مقایسه آن با شبکه خود رمزنگار کلاسیک در حالت شرایط نویزی و با نسبت سیگنال به نویز ۳dB
Table 2. Investigating the performance of the variable self-encryptor network in reducing input dimensions and comparing it with the classic self-encryptor network in noisy conditions and with a signal-to-noise ratio of 3dB.

Ref.	Feature Reduction	Category Function	Precision	Sensitivity	Being Special
-----	All features	KNN	62.88	62.76	62.33
		SVM	63.78	64.55	63.87
Proposed method	Classic self-encryption	KNN	73.88	75.12	method
		SVM	78.35	78.29	79.12
Proposed method	Variable self-encryption	KNN	82.86	81.33	82.97
		SVM	85.91	82.32	83.42

جدول ۲ نشان می‌دهد که با اضافه‌شدن سیگنال تنفسی به سیگنال PCG، عملکرد طبقه‌بندی کاهش می‌یابد. بااین‌حال، روش مبتنی بر خود رمزنگار متغیر عملکرد بهتری را در مقایسه با سایر خود رمزنگار کلاسیک نشان می‌دهد و این برتری به طور محسوس قابل‌مشاهده است. این توانایی، طرح پیشنهادی را برای طبقه‌بندی سیگنال‌های PCG در میان نویز تنفسی بالا نشان می‌دهد. در نهایت عملکرد رویکرد پیشنهادی با تکنیک‌های موجود مقایسه شده است. در جدول ۳، مقایسه کلی روش پیشنهادی با هفت تکنیک به‌روز نشان‌داده‌شده است. باید به این نکته اشاره کرد به‌منظور مقایسه منصفانه، مراجع با پایگاه‌داده مشابه با کار خودمان برای مقایسه انتخاب شده است.

جدول ۳: مقایسه عملکرد روش پیشنهادی با جدیدترین روش¬های موجود
Table 3. Comparing the performance of the proposed method with the latest available methods.

Ref.	Work method	Precision	Sensitivity	Being Special
[12]	Feature extraction using MFCC and MFSC, and using deep networks	97.10	99.26	94.86
[14]	Feature extraction using MFCC and a convolutional network	84.15	80.63	87.66
[18]	Feature extraction from the time domain and a deep neural network	85.65	86.73	84.75
[20]	Using Wavelet Transform and Convolutional Neural Network	98	98	98
[31]	Feature extraction using MFCC and short-time Fourier transform, using convolutional neural networks and LSTM neural networks	91.31	93.80	92.60
[32]	Combining time, frequency, and time-frequency domain features and employing deep neural networks	94.5	98.2	92.60
[3]	Cochleogram transformation and application of deep neural networks	98.33	98.20	98.45
Proposed method	Cochleogram transformation and application of deep neural networks	99.85	99.95	99.42

۴- نتیجه‌گیری

در این مقاله به‌کارگیری شبکه خود رمزنگار متغیر برای استخراج ویژگی از بردار به‌دست‌آمده از تبدیل کوچلیگرام را برای طبقه‌بندی صدای قلب مورد مطالعه قرار داده‌ایم. یک مجموعه‌داده متعادل از سیگنال PCG نرمال و غیرنرمال شامل ۳۴۸۲ سیکل صدای قلب برای آموزش و ارزیابی روش پیشنهادی استفاده شده است. بانک فیلتر گاماتون برای هر فریم از سیگنال، یک بردار ۶۴ بعدی که معادل پاسخ فرکانسی سیستم شنوایی انسان است را تولید می‌کند. این فیلتر بانک برای استخراج ویژگی

کوچلیگرام استفاده شده است و ثابت شده است که در پردازش صدای قلب بسیار دقیق و کارآمد است. در ادامه به کمک طراحی یک شبکه خود رمزنگار متغیر و با دسترسی به لایه پنهان این شبکه، ۱۰ ویژگی مفید استخراج می‌شود. به بیان دیگر از شبکه خود رمزنگار برای کاهش ابعاد استفاده شده است. یافته‌های عددی، اثر بخشی رویکرد پیشنهادی برای دسته‌بندی صدای قلب را اثبات می‌کنند به طوری که با استفاده از تابع دسته‌بند SVM به‌دقت ۹۹/۵۵، حساسیت ۹۸/۷۵ و ویژه بودن ۹۹/۷۰ دست یافته‌ایم. این مدل می‌تواند به‌عنوان یک ابزار تشخیصی برای تشخیص ناهنجاری‌های قلبی به پزشکان کمک کند. برای توسعه یک سیستم تشخیص بیماری قلبی مبتنی بر صدای قلب، این مدل می‌تواند توسط یک رایانه شخصی آموزش داده شود و به‌عنوان یک سیستم قابل‌حمل برای مصارف بالینی استفاده شود.

مراجع

[1] Cardiovascular diseases (CVDs): WHO Fact Sheet. WHO (11 June 2021). https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).

[2] E. Etchells, C. Bell, and K. Robb, “Does this patient have an abnormal systolic murmur?,” *Jama*, vol. 277, no 7, pp. 564–571, 1997.

[3] S. Das, S. Pal, and M. Mitra, “Deep learning approach of murmur detection using Cochleagram,” *Biomedical Signal Processing and Control*, vol. 1, no. 77, p. 103747, 2022, doi: 10.1016/j.bspc.2022.103747.

[4] S.K. Randhawa and M. Singh, “Classification of heart sound signals using multi-modal features,” *Procedia Computer Science*, vol. 58, pp. 165-171, 2015, doi: 10.1016/j.procs.2015.08.045.

[5] F. Ghaderi, H. R. Mohseni, and S. Sanei, "Localizing Heart Sounds in Respiratory Signals Using Singular Spectrum Analysis," in *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 12, pp. 3360-3367, Dec. 2011, doi: 10.1109/TBME.2011.2162728.

[6] Y. Zeinali and S.T.A. Niaki, “Heart sound classification using signal processing and machine learning algorithms,” *Machine Learning with Applications*, vol. 7, p. 100206. 2022, doi: 10.1016/j.mlwa.2021.100206.

[7] F. Safara and A.R.A. Ramaiah, “RenyiBS: Renyi entropy basis selection from wavelet packet decomposition tree for phonocardiogram classification,” *The Journal of Supercomputing*, vol. 77, pp. 3710-3726, 2021, doi: 10.1007/s11227-020-03413-9.

[8] B. Daoud, K. Nayad, B. Braham, and B. Messaoud, “Heart murmurs detection and characterization using wavelet analysis with Renyi entropy,” *Journal of Mechanics in Medicine and Biology*, vol. 17, no. 6, 2017, doi: 10.1142/S0219519417500932.

[9] E. Kay and A. Agarwal, "DropConnected neural network trained with diverse features for classifying heart sounds," *Computing in Cardiology Conference (CinC)*, Vancouver, BC, Canada, 2016, pp. 617-620.

[10] G. Eslamizadeh and R. Barati, “Heart murmur detection based on wavelet transformation and a synergy between artifcial neural network and modified neighbor annealing methods,” *Artif. Intell. Med.*, vol. 78, pp. 23-40, 2017, doi: 10.1016/j.artmed.2017.05.005.

[11] H. Li, Y. Ren, G. Zhang, R. Wang, J. Cui, and W. Zhang, "Detection and Classification of Abnormities of First Heart Sound Using Empirical Wavelet Transform," in *IEEE Access*, vol. 7, pp. 139643-139652, 2019, doi: 10.1109/ACCESS.2019.2943705.

[12] H. Chowdhury, K. N. Poudel, and Y. Hu, "Time-Frequency Analysis, Denoising, Compression, Segmentation, and Classification of PCG Signals," in *IEEE Access*, vol. 8, pp. 160882-160890, 2020, doi: 10.1109/ACCESS.2020.3020806.

[13] J. Chen, Y. Wang, and D. Wang, "A Feature Study for Classification-Based Speech Separation at Low

Signal-to-Noise Ratios," in *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 12, pp. 1993-2002, Dec. 2014, doi: 10.1109/TASLP.2014.2359159.

[14] V. Maknickas and A. Maknickas, “Recognition of normal–abnormal phonocardiographic signals using deep convolutional neural networks and mel-frequency spectral coeffcients,” *Physiol. Meas*., vol. 38, no. 8, pp. 1671-1684, 2017, doi: 10.1088/1361-6579/aa7841.

[15] M. Hamidi, H. Ghassemian, and M. Imani, “Classifcation of heart sound signal using curve ftting and fractal dimension,” *Biomed. Signal Process. Control*, vol. 39, pp. 351–359, 2018, doi: 10.1016/j.bspc.2017.08.002.

[16] M.S. Ahmad, J. Mir, M.O. Ullah, M.L.U.R. Shahid, and M.A. Syed, “An effcient heart murmur recognition and cardiovascular disorders classifcation system,” *Australas. Phys. Eng. Sci. Med*., vol. 42, no. 3,pp. 733–743, 2019, doi: 10.1007/s13246-019-00778-x.

[17] G. Yaseen, Y. Son, and S. Kwon, “Classifcation of heart sound signal using multiple features,” *Appl. Sci*., vol. 8, no. 12, 2018, doi: 10.3390/app8122344.

[18] P.T. Krishnan, P. Balasubramanian, and S. Umapathy, “Automated heart sound classifcation system from unsegmented phonocardiogram (PCG) using deep neural network,” *Phys. Eng. Sci. Med*., vol. 43, no. 2, pp. 505–515,2020, doi: 10.1007/s13246-020-00851-w.

[19] S.L. Oh, V. Jahmunah, C.P. Ooi, R.-S. Tan, E.J. Ciaccio, T. Yamakawa, M. Tanabe, M. Kobayashi, and U. Rajendra Acharya, “Classifcation of heart sound signals using a novel deep WaveNet model,” *Comput. Methods Programs Biomed*., vol. 196, p. 105604, 2020, doi: 10.1016/j.cmpb.2020.105604.

[20] P. Dhar, S. Dutta, and V. Mukherjee, “Cross-wavelet assisted convolution neural network (AlexNet) approach for phonocardiogram signals classifcation,” *Biomed. Signal Process. Control.,* vol. 63, p. 102142, 2021, doi: 10.1016/j.bspc.2020.102142.

[21] J. P. Dominguez-Morales, A. F. Jimenez-Fernandez, M. J. Dominguez-Morales, and G. Jimenez-Moreno, "Deep Neural Networks for the Recognition and Classification of Heart Murmurs Using Neuromorphic Auditory Sensors," in *IEEE Transactions on Biomedical Circuits and Systems*, vol. 12, no. 1, pp. 24-34, Feb. 2018, doi: 10.1109/TBCAS.2017.2751545.

[22] S. Das, S. Pal, and M. Mitra, “Acoustic feature based unsupervised approach of heart sound event detection,” *Comput Biol Med.,* vol. 126, p. 103990, 2020, doi: 10.1016/j.compbiomed.2020.103990.

[23] C. Liu *et al.*, “An open access database for the evaluation of heart sound algorithms,” *Physiol. Meas*., vol. 37, no. 12, pp. 2181–2213, 2016, doi: 10.1088/0967-3334/37/12/2181.

[24] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J. E. Mietus, G.B. Moody, C.-K. Peng, and H.E. Stanley, “PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals,” *Circulation*, vol. 101, no. 23, 2000, pp. 215-220, doi: 10.1161/01.cir.101.23.e215.

[25] C. Xu, S.J. Chao, G. Jiang, X. Chen, Q. He, and P. Xie, “Two-level multidomain feature extraction on sparse representation for motor imagery classifcation,” *Biomedical Signal Processing and Control*, vol. 62, p. 102160, 2020, doi: 10.1016/j.bspc.2020.102160.

[26] H. Chang and J. Yang, “Genetic-based feature selection for efcient motion imaging of a brain–computer interface framework,” *Journal of neural engineering, vol.* 15, no. 5, 2018, doi: 10.1088/1741-2552/aad567.

[27] P. Tan, X. Wang, and Y. Wang, “Dimensionality reduction in evolutionary algorithms-based feature selection for motor imagery braincomputer interface,” *Swarm and Evolutionary Computation*, vol. 52, p. 100597, 2020, doi: 10.1016/j.swevo.2019.100597.

[28] Y. Baysal, S. Ketenci, I.H. Altas, and T. Kayikcioglu, “Multi-objective symbiotic organism search algorithm for optimal feature selection in brain computer interfaces,” *Expert Systems with Applications*,

vol. 165, no. 113907, 2021, doi: 10.1016/j.eswa.2020.113907.

[29] M. A. Nadoomi and S. Majid, "Using Ant Colony Algorithm and Pairwise Learning to Classify Attack in Intrusion Detection Systems," *Journal of Communication Engineering*, vol. 9, no. 36, pp. 39–53, 2020 [in Persian].

[30] I. Jamali, S. Javad Mirabedini, and A. Haronabadi, "Offering a Model for Persian Texts Classify by Combination of Classification Methods," *Journal of Communication Engineering,* vol. 10, no. 38, pp. 61-72, 2020 [in Persian].

[31] F. Hajiani, N. Parhizgar, and A. Keshavarz, "Hyperspectral Image Classification Using Low Rank Representation and Spectral-Spatial Information," *Journal of Southern Communication Engineering*, vol. 11, no. 43, pp. 27-38, 2022, doi: 10.30495/jce.2022.689206 [in Persian].

[32] M. Jalali and T. Sedghi, "Extraction of Multiple Hybrid Features to Reduce the Semantic Vacuum with the Semi-Supervised Classification," *Journal of Southern Communication Engineering*, vol. 12, no. 45, pp. 31-44, 2022, doi: 10.30495/jce.2022.691134 [in Persian].

[33] Innocent Tatchum Sado, Louis Fippo Fitime, Geraud Fokou Pelap, Claude Tinku, Gaelle Mireille Meudje, Thomas Bouetou Bouetou, "Early multi-cancer detection through deep learning: An anomaly detection approach using Variational Autoencoder, " *Journal of Biomedical Informatics*, Vol 160, 2024, no. 104751, doi: 10.1016/j.jbi.2024.104751.

JOURNAL OF SOUTHERN COMMUNICATION ENGINEERING
ISLAMIC AZAD UNIVERSITY BUSHEHR BRANCH

E-ISSN: 2980-9231
https://sanad.iau.ir/journal/jce

https://doi.org/10.30495/jce.2025.1993480.1336

Vol. 14/ No. 56/2025

Research Article

Presenting an Attack-Resistant Communication Model for Secure Routing in Underwater Sensor Networks

Tayebeh Nourali Ahari [1] | Mehdi Sadeghzadeh [2]

[1]Department of Information Technology Management, SR.C., Islamic Azad University, Tehran, Iran.
tayebeh.ahari@srbiau.ac.ir

[2]Department of Computer Engineering, SR.C., Islamic Azad University, Tehran, Iran
mehdi.sadeghzadeh@iau.ac.ir

Correspondence
Mehdi Sadeghzadeh, Associate Professor of Computer Engineering, SR.C., Islamic Azad University, Tehran, Iran.
mehdi.sadeghzadeh@iau.ac.ir

Main Subjects:
Secure Routing in Underwater Sensor Networks
Paper History:
Received: 14 April 2024
Revised: 17 August 2024
Accepted: 18 August 2024

Abstract

Underwater communication networks face numerous challenges, including routing, signal interference, energy consumption, and security threats. Although routing protocols are optimized for resilience against common disturbances in underwater environments, they are not specifically designed to counter attacks or malicious neighbor nodes. Key security threat factors in underwater sensor networks include limited power sources, constrained communication media, and harsh underwater conditions. Therefore, this research aims to develop a secure communication model resistant to routing attacks in underwater sensor networks. Two communication link models were considered: Scenario 1 uses a basic distance-based model, while Scenario 2 employs a probabilistic channel gain model between node pairs. Simulation results encompass four stages: 1) secure neighbor discovery under wormhole attacks; 2) initial route discovery and selection of reliable nodes for data forwarding to the sink; 3) attack detection during data distribution based on node status information for identifying Sybil attacks; 4) alternative secure path discovery for detecting malicious nodes. The proposed scheme demonstrated higher success rates compared to the basic approach and lower mobility energy costs, achieving comparable success performance.

Keywords: Underwater Internet of Things, Safe Neighbor Discovery, Underwater Routing, Wormhole Attack.

Highlights

- Limitation of processing and communication capabilities in underwater Internet of Things.
- Simulation of guide signal transmission and neighbor table formation in two communication models.
- The proposed method demonstrated the highest network throughput compared to the basic method.

Citation: T. Nourali Ahari, and M. Sadeghzadeh, "Presenting an Attack-Resistant Communication Model for Secure Routing in Underwater Sensor Networks," *Journal of Southern Communication Engineering*, vol. 14, no. 56, pp. 58–74, 2025, doi:10.30495/jce.2025.1993480.1336 [in Persian].

1. Introduction

The introduction discusses the importance of neighbor discovery in underwater IoT networks and the risks posed if malicious nodes infiltrate the process. It proposes using physical layer authentication with trusted and sink nodes to ensure data security. This method leverages the unique characteristics of underwater acoustic channels for authentication. The approach is validated through simulations and marine experiments.

The paper also addresses privacy challenges in IoT networks, emphasizing the need for robust solutions. The article is structured as follows: Section 2 covers network modeling, Section 3 reviews the simulation steps, and Section 4 presents the results, with a concluding discussion in the final section.

2. Innovation and Contributions

This paper addresses key challenges in underwater IoT networks through a novel physical layer authentication method. Key innovations include:

- **Distributed authentication** using trusted nodes and sink nodes to enhance security.
- **Channel-based verification** leveraging spatial and temporal characteristics of underwater acoustic channels for reliable message authentication.
- **Comprehensive simulation** of the method in an underwater IoT voice communication network to assess wormhole attack impacts.
- **Experimental validation** through extensive numerical simulations and real-world marine tests.
- **Robust privacy solutions** for protecting location, device, and data privacy in distributed IoT environments.

These advancements strengthen data confidentiality, integrity, and availability in underwater IoT networks, enabling secure communication in challenging underwater conditions.

3. Materials and Methods

NS-2 is utilized for software simulation, capitalizing on its comprehensive library of network protocols. The simulation comprises four key stages:

1. Secure neighbor discovery under wormhole attack conditions
2. Initial path discovery to enhance packet delivery reliability
3. Attack detection during data distribution
4. Identification of alternative safe paths to detect malicious nodes

The neighbor discovery process involves broadcasting digitally signed request messages containing unique node identifiers.

4. Results and Discussion

The simulation results were obtained by averaging 40 runs for each output. Key findings include:

- **Packet Delivery Rate**: The proposed method maintains a high success ratio (above 93%) for networks of up to 400 nodes, outperforming the baseline method in both success rate and mobility energy cost.
- **Network Throughput**: Consistently higher throughput is achieved across all simulations, demonstrating superior performance to baseline methods even with increased hop counts.
- **End-to-End Delay**: While average delay generally decreases with additional nodes, a slight increase is observed at very high node densities. Overall, delays remain lower than in previous studies.
- **Energy Consumption**: The method exhibits improved energy efficiency, particularly in high-density networks, attributable to optimized path selection for data transmission.

These results demonstrate the effectiveness of the proposed approach in delivering reliable, efficient, and secure performance for underwater IoT networks.

5. Conclusion

This paper presents a simulation study of neighbor and path discovery mechanisms in underwater acoustic communication networks, specifically designed for IoT applications with constrained processing and communication capabilities. The investigation examines beacon signal transmission and neighbor table formation using two distinct communication models while evaluating the performance degradation caused by malicious nodes.

Key findings indicate that the presence of malicious nodes has a significant impact on network operations. Comparative analysis with prior studies reveals superior performance across multiple metrics:

- Throughput efficiency
- Reduced end-to-end delay
- Enhanced packet delivery rate (achieving 94%)
- Optimized energy consumption

The proposed method shows consistent improvement over existing approaches in all measured performance indicators, establishing its effectiveness for resource-constrained underwater IoT deployments.

6. Acknowledgement
There was no financial support for this article. But we are grateful for the valuable contributions made by the reviewers in shaping this paper into its final form.

Declaration of Competing Interest: The Authors do not have a conflict of interest. The authors approve the content of the paper.

Author Contributions:
Mehdi Sadeghzadeh: verified the analytical methods; **Tayebeh Nourali Ahari**: made the simulations and wrote the manuscript.

E-ISSN: 2980-9231
https://sanad.iau.ir/journal/jce

مهندسی مخابرات جنوب
دانشگاه آزاد اسلامی واحد بوشهر

https://doi.org/10.30495/jce.2025.1993480.1336

مقاله پژوهشی

ارائه یک مدل ارتباطی مقاوم در برابر حملات به‌منظور مسیریابی امن در شبکه‌های حسگر زیر آب

طیبه نورعلی آهاری[۱] | مهدی صادق زاده*[۲]

[۱] گروه مدیریت فناوری اطلاعات، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران.
tayebeh.ahari@srbiau.ac.ir

[۲] گروه مهندسی کامپیوتر، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران.
mehdi.sadeghzadeh@iau.ac.ir

نویسنده مسئول
*مهدی صادق زاده، دانشیار گروه علمی مهندسی کامپیوتر، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران.
mehdi.sadeghzadeh@iau.ac.ir

موضوع اصلی:
مسیر یابی امن در شبکه های حسگر زیر آب

تاریخچه مقاله:
تاریخ دریافت: ۲۶ فروردین ۱۴۰۳
تاریخ بازنگری: ۲۷ مرداد ۱۴۰۳
تاریخ پذیرش: ۲۸ مرداد ۱۴۰۳

چکیده:

شبکه‌های ارتباط زیرآب با چالش‌هایی مانند مسیریابی، تداخل سیگنال، مصرف انرژی و تهدیدات امنیتی مواجه هستند. اگر چه پروتکل‌های مسیریابی برای محیط‌های زیرآبی به‌منظور مقاومت در برابر اختلالات معمول بهینه شده‌اند، اما برای مقابله با حملات و رفتارهای مخرب گره‌های همسایه به‌طور خاص طراحی نشده‌اند. عوامل اصلی تهدیدات امنیتی در شبکه‌های حسگر زیرآب شامل منبع تغذیه محدود، رسانه‌های ارتباطی محدود و شرایط دشوار ارتباط زیرآب هستند. بنابراین هدف از این پژوهش ارائه مدل ارتباطی مقاوم در برابر حملات مسیریابی امن در شبکه‌های حسگر زیر آب است. بدین منظور دو مدل برای پیوندهای ارتباطی بین هر جفت گره در نظر گرفته شد. سناریو ۱، یک مدل پایه مبتنی بر فاصله است. در سناریوی دوم، از یک مدل احتمالی از بهره کانال بین هر جفت گره استفاده گردید. نتایج شبیه‌سازی شامل چهار مرحله: ۱) کشف همسایه امن تحت حملات کرم چاله ۲) فرآیند کشف مسیر اولیه و انتخاب گره های قابل اعتماد برای انتقال به سینک ۳) فرآیند تشخیص حمله در حین توزیع داده بر اساس اطلاعات وضعیت گره‌ها برای شناسایی حمله سیبل و ۴) کشف مسیر امن جایگزین برای شناسایی گره‌های مخرب می‌باشد، نشان داد که طرح پیشنهادی به نرخ موفقیت بهتری نسبت به طرح پایه دست یافته و با هزینه‌ی انرژی تحرک کمتری نسبت به روش پایه ، نسبت موفقیت قابل‌مقایسه‌ای را به دست آورده است.

کلید واژه‌ها: اینترنت اشیای زیر آب، کشف همسایه امن، مسیریابی زیر آب، حمله کرم چاله

تازه‌های تحقیق:

- محدودیت قابلیت‌های پردازش و ارتباطی در اینترنت اشیا زیر آب.
- شبیه‌سازی انتقال سیگنال راهنما و شکل‌گیری جدول همسایه در دو مدل ارتباطی.
- روش پیشنهادی بالاترین توان عملیاتی شبکه را در مقایسه با روش پایه نشان داد.

۱-مقدمه

اینترنت اشیا به‌عنوان جزئی از انقلاب صنعتی سوم که شی محاسباتی را برای ارسال و دریافت داده‌های فیزیکی از طریق اینترنت جاسازی می‌کند، مورد استقبال زیادی قرار گرفته است [۱]. اینترنت اشیا می‌تواند به‌صورت تلفیقی از یک شبکه ناهمگن در نظر گرفته شود که نه‌تنها چالش‌های امنیتی در شبکه‌های حسگر فعلی، ارتباطات مخابراتی تلفن همراه و اینترنت را در بردارد بلکه مسائلی همچون شبکه خصوصی، احراز هویت در شبکه ناهمگن، چالش‌های کنترل دسترسی و مسیریابی امن در بین دستگاه‌های ناهمگن را نشان می‌دهد [۲].

شبکه اینترنت اشیای زیر آب، متشکل از وسایل نقلیه خودران، حسگرهای از راه دور، رله‌های سطحی، و نظایر آن، به دلیل توانایی‌شان در پایش و بررسی مناطق اقیانوسی بزرگ، در سال‌های اخیر مورد توجه قرار گرفته است. شبکه‌های ارتباطی زیر آب با مجموعه چالش‌های متفاوتی نسبت به شبکه‌های ارتباطی زمینی، از جمله مسیریابی بسته، اختلال سیگنال، مصرف انرژی و حملات احتمالی در شبکه، مواجه است. از این‌رو در مطالعات متعددی به موضوع پروتکل‌های مسیریابی که شرایط منحصربه‌فرد تجربه‌شده در محیط زیر آب را کنترل می‌کنند[۳]، [۴]، پرداخته شده است. برای مثال، دنیل جی و همکارانش در مرکز امنیت ملی هیوم [۵] یک مدل شبیه‌سازی کشف همسایه و مسیر در شبکه ارتباطات صوتی زیر آب را ارائه کردند که در آن بر کاربرد اینترنت اشیای زیر آب که در آن توانمندی‌های پردازشی برای رمزگذاری/هویت‌سنجی و توانمندی‌های ارتباطی برای تصدیق یا بازانتقال را ارائه داده‌اند که ممکن است محدود باشند. آنها در مدل شبیه‌سازی‌شان تأثیر کانال زیر آب و رفتار گره مخرب بر کشف مبتنی بر سیگنال راهنما را اندازه‌گیری کردند. نتایج این بررسی نشان داد که گره مخرب می‌تواند با موفقیت بر عملکرد شبکه تأثیر بگذارد، با این حال، تأثیر کانال از اهمیت بیشتری برخوردار بود. بنابراین آنها نتیجه‌گیری کردند که، مسیریابی چندمسیره استراتژی کاهشی بالقوه‌ای هم برای شرایط کانال و هم برای حمله فروچاله خواهد بود.

قریشی[۱] و همکارانش [۶] پروتکل مسیریابی فرصت‌طلبانه اجتناب از حفره[۲] (OVAR) را برای شبکه‌های زیر آب مطرح می‌کند که به مسئله حفره‌هایی می‌پردازد که ممکن است طرح‌های مسیریابی رایج‌تر را مختل کنند. با اینکه پروتکل‌هایی همچون OVAR در مقابل اختلالات معمول تجربه‌شده در محیط زیر آب تاب‌آوری دارند، لزوماً برای کنترل حملات احتمالی و رفتار مخرب گره‌های همسایه طراحی نشده‌اند. آنها در مطالعه خود بر اهمیت قابلیت مقیاس‌پذیری در شبکه‌های زیر آبی با تعداد زیاد گره‌ها تأکید می‌کنند.

در بحث چالش‌های شبکه‌های حسگر بیسیم زیرآب می‌توان به این نکات اشاره کرد که در کانال‌های صوتی که پهنای باند کم و کیفیت لینک ضعیفی دارند، به دلیل محوشدگی و ویژگی‌های انعکاسی کانال صوتی، نرخ خطای بیت[۳] بالاست. با وجود این، ارتباطات صوتی بازه انتقال بهتری نسبت به فرکانس رادیویی/ القای مغناطیسی[۴] (MI) با پهنای باند بالاتر یا انتقال نوری زیر آب دارند [۷].

با توجه به سیار بودن حسگرها، عدم قطعیت افزوده در کشف و نگهداری همسایه‌ها را باید در نظر گرفت. گره‌ها در محیطی با عمق و با جریان‌های آبی متغیری قرار دارند که عدم قطعیت بیشتری در اتصال و مکان‌های آینده گره‌ایجاد می‌کنند.

شبکه‌های زیر آب در معرض انواع گوناگونی از حملات قرار دارند از این‌رو امنیت یکی مهم‌ترین نگرانی‌ها در شبکه‌های زیر آب وسایل نقلیه خودران و حسگرهاست [۸]، [۷]، از جمله حملات کرم‌چاله،حملات فروچاله[۵]، ارسال انتخابی[۶] و بسیاری موارد دیگر[۶]، [۹]، [۸]، [۱۰].برای مثال، در حملات کرم چاله، بسته‌های یک منطقه از شبکه از طریق لینک سریع و خارج از باند، به منطقه دیگری از شبکه منتقل شده و بازپخش می‌شوند، به‌طورکلی این حملات، با نامطمئن کردن اطلاعات مربوط به همسایه‌ها، پروتکل کشف مسیر شبکه را هدف قرار می‌دهند. فروچاله این کار را با همه‌پخشی[۷] شماره هاپ پایین یا اولویت مسیریابی بالا انجام می‌دهد و سپس همه بسته‌های دریافتی‌اش را حذف می‌کند. حمله کرم‌چاله دو گره دشمن را با کانال

---

[1] Ghoreishi
[2] opportunistic void avoidance routing
[3] bit error rate
[4] magnetic induction
[5] sinkhole attack
[6] selective forwarding
[7] broadcasting

ارتباطی سریع (مثلاً فرکانس رادیویی زمینی) به هم وصل می‌کند تا با استفاده از آن به‌سرعت اطلاعات سیگنال راهنما را ارسال کنند. آنها این اطلاعات را منتقل می‌کنند و شبکه نتیجه می‌گیرد که این دو گره به هم نزدیک‌اند، در نتیجه به تمام گره‌های مجاور اطلاعات مسیریابی نادرست می‌دهد. گره‌های مهاجم همچنین می‌توانند شبکه را متقاعد کنند که گره یکسانی در دو مکان متفاوت‌اند. در برخی مطالعات اخیر به این چالش‌ها پرداخته شده است. برای مثال؛ درگاهی و همکاران در مطالعه‌ای به بررسی چالش‌های امنیتی و حملات به شبکه‌های حسگر بیسیم زیر آب و همچنین راه‌حل‌های احتمالی کاهش این چالش احتمالی می‌پردازند [۱۱]. در برخی پژوهش‌ها نیز پروتکل‌های مسیریابی به‌طور خاص برای محیط‌های ارتباطی زیر آب طراحی شده‌اند [۵]، [۶]، [۱۲]-[۱۳]. به‌ویژه، محققان پروتکل‌هایی همانند OVAR [۶]، IVAR [۱۴]، و BEAR [۱۳] را ارائه کرده‌اند. پروتکل مسیریابی فرصت‌طلبانه اجتناب از حفره، یا OVAR، سعی می‌کند که توان عملیاتی را در محیط پراکنده و پراتلاف زیر آب، بالا ببرد و در عین حال در انرژی صرفه‌جویی کند [۶]. مسیریابی ٔ اجتناب از حفره[۱]، یا IVAR، مشابه پروتکل اجتناب از حفره‌ای است که سعی می‌کند از مشخص بودن مقصد برای داده‌ها به نفع خودش استفاده کند و از مکان مقصد و شماره هاپ گره‌های انتقال‌دهنده برای انتخاب انتقال بعدی بهره ببرد [۱۴]. طرح مسیریابی سازگار با انرژی متوازن[۲] (BEAR) به دنبال طولانی کردن طول عمر حسگرهای زیر آب از طریق ارزیابی طرح‌های مصرف انرژی کارآمد و مصرف انرژی متوازن است. پروتکل مسیریابی فشار به حفره (VAPR) از سونوبوی‌ها[۳] برای دور زدن مناطق حفره‌دار و انتقال داده به سینک استفاده می‌کند [۱۵]. در همین راستا ژانگ[۴] و همکارانش در مطالعه‌شان چند پروتکل کشف مسیر را ارائه کردند، آنها اذعان داشتند که مصرف بالای انرژی، ساختار پویا و تاخیر انتقال، نیازمند پروتکل‌های مسیریابی ویژه‌ای هستند، که عبارتند از:

- پروتکل‌های مبتنی بر انرژی: این پروتکل‌ها بر بهینه‌سازی مصرف انرژی در گره‌های حسگر زیرآبی تمرکز دارند.
- پروتکل‌های مبتنی بر داده: این پروتکل‌ها به رویکردهای مرتبط با داده برای انتقال کارآمد تمرکز دارند.
- پروتکل‌های مبتنی بر اطلاعات جغرافیایی: این پروتکل‌ها از اطلاعات مکانی برای تصمیم‌گیری‌های مسیریابی استفاده می‌کنند.

همچنین عواملی مانند تداخل نویز محیطی و تغییر فرکانس داپلر تأثیری بر کیفیت ارتباط گره‌های حسگر زیرآبی دارند [۱۶]. با اینکه رمزنگاری راه‌حل رایجی برای حفظ این فرایند از حملاتی چون فروچاله و کرم‌چاله است، ممکن است راه‌حلی به‌شدت پردازشی باشد که دستگاه‌های اینترنت اشیای زیر آب با توان محدود قادر به پشتیبانی از آن نباشند.

تخمین مسیر رسیدن[۵] (DoA) مفید است، چون با دانستن اینکه سیگنال از کجا می‌آید، الگوریتم‌های شکل‌دهی پرتو تطبیقی[۶] می‌توانند توان سیگنال‌های تداخلی را به حداقل برسانند. تخمین DoA معمولاً به آرایه‌ای از حسگرهای برداری یا آرایه‌ای از هیدروفون‌ها[۷] نیاز دارد که نیروهای آکوستیک صفحه‌ای[۸] را تشخیص می‌دهند و بعد اختلاف فاز سیگنال‌های دریافتی را برای تخمین جهت موردنظر تحلیل می‌کنند. اطلاعات DoA همچنین راهی برای شناسایی بالقوه گره‌هایی فراهم می‌کند که مطابق با مکان جغرافیایی‌شان رفتار نمی‌کنند (یا اطلاعات را گزارش نمی‌دهند).

با توجه به مطالبی که عنوان شد، و مطابق با نتایج مطالعات پیشین ایمن‌سازی به دلیل ویژگی‌ها و محدودیت‌های آنها آسان نیست و همچنین استقرار و نگهداری شبکه از سوی دیگر هزینه بالائی در برداشته و طرح‌ها از نظر مصرف انرژی بار سنگینی خواهند داشت.

همچنین با توجه به پیش‌بینی‌ناپذیر بودن محیط زیر آب، کشف همسایه نقش مهمی در تعیین مسیرهای دقیق انتقال بسته در شبکه زیر آب ایفا می‌کند. متأسفانه، اگر اقدامات احتیاطی درستی اتخاذ نشود، به‌راحتی می‌توان از کشف همسایه سوءاستفاده کرد. اگر گره مخرب بتواند به فرایند کشف همسایه نفوذ کند و به گره داخلی تبدیل شود، می‌تواند به داده‌های حساس گوش بدهد. برای جلوگیری از این اتفاق، گره‌ها باید با پروتکل‌هایی برنامه‌ریزی شوند که آنها را قادر به شناسایی گره‌هایی می‌کنند که، مثلاً براساس رفتار آموخته‌شده، عضو قانونی شبکه نیستند. در این مقاله، شبکه‌ای از دستگاه‌های اینترنت اشیای زیر آب را

[1] inherently void avoidance routing
[2] balanced energy adaptive routing
[3] sonobuoys
[4] R. Zhang
[5] DoA - Direction of arrival.
[6] adaptive beamforming algorithms
[7] hydrophones
[8] plane acoustics

شبیه‌سازی خواهیم کرد که در محیط ارتباطی صوتی فعالیت دارند و تأثیر حمله کرم چاله در چنین شبکه‌ای را بررسی می‌کنیم. از چارچوب شبیه‌سازی بهره می‌بریم تا امکان تصویرسازی فرایند کشف شبکه فراهم شود. از دیگر سو هدف این مقاله، بررسی آسیب‌پذیری طرح‌های مسیریابی زیر آب در مقابل رفتار مخرب در فرایند کشف مسیر شبکه می‌باشد. در واقع شبیه‌ساز شبکه ارتباطی زیر آب پویایی در این مطالعه ایجاد خواهد شد که روال سیگنال‌دهی را پیاده‌سازی کرده که می‌توان آن را اصلاح کرد تا طرح‌های مسیریابی پیشنهاد شده در منابع مطالعاتی را نشان دهد [۱۷].

رویکرد در نظر گرفته شده برای افزایش محرمانگی، یکپارچگی و در دسترس بودن داده‌ها احراز هویت لایه فیزیکی است که یک روش مشارکتی برای احراز هویت پیام‌ها در شبکه‌های صوتی زیر آب معرفی شده است .این شامل گره‌های قابل اعتماد و گره سینک است که برای تعیین اینکه آیا پیام دریافتی قانونی است یا از طرف یک مهاجم با یکدیگر همکاری می‌کنند .این روش از وابستگی مکانی و عدم تغییر زمانی ویژگی‌های کانال صوتی زیر آب برای محاسبه یک شاخص تصمیم برای احراز هویت استفاده می‌کند .گره‌های مورد اعتماد محاسبه را به‌صورت توزیع شده انجام می‌دهند، درحالی‌که گره سینک نظرات آنها را ترکیب می‌کند و بدون نیاز به ارائه بازخورد به گره‌های مورد اعتماد تصمیم نهایی را می‌گیرد .اثربخشی این رویکرد در شبیه‌سازی‌های عددی گسترده و آزمایش‌های دریایی در مطالعات دیگر تائید شده است. در محیط‌هایIoUT ، گره‌های حسگر به‌صورت پراکنده مستقر هستند، که مدیریت حریم خصوصی را بسیار دشوار می‌کند .بنابراین این محیط‌ها حساس و پیچیده در نظر گرفته می‌شوند .در زمینهIoUT ، اصطلاح "حریم خصوصی" معنای گسترده‌تری دارد و مکان، دستگاه و حریم خصوصی داده‌ها را پوشش می‌دهد . ازاین‌رو، راه‌حل‌های قوی برای حفظ حریم خصوصی در شبکه‌های IoUT ضروری است .رویکردهای FL به رفع نگرانی‌های مربوط به اعتماد کمک می‌کند .اما FL دارای محدودیت‌های حریم خصوصی خاصی است .پرداختن به مشکلات مربوط به حریم خصوصی در تنظیمات دارای حریم خصوصی یک چالش تحقیقاتی باز است. پیاده‌سازی پروتکل‌های سبک وزن و ایمن می‌تواند یک راه حل ممکن باشد [۱۸].

ساختار مقاله به این صورت است که سازمان‌دهی شده که در بخش ۲ به توضیح نحوه مدل‌سازی و شبیه‌سازی شبکه ارتباطی زیرآب، از جمله مدل لینک ارتباطی و مدل انرژی پرداخته خواهد شد. در بخش ۳ مراحل شبیه‌سازی به اجمال بررسی‌شده و در ادامه در بخش ۴ نتایج شبیه‌سازی ارائه می‌گردد، در نهایت در بخش نتیجه‌گیری به بحث و بررسی حول نتایج به دست آمده در شبیه‌سازی خواهیم پرداخت.

۲-مدل شبکه

داده‌ها و ویژگی‌های پژوهش پیش رو از داده‌های پژوهش‌های پیشین اقتباس شده است. در این بخش، نحوه مدل‌سازی و شبیه‌سازی شبکه ارتباطی زیر آب، از جمله مدل لینک زیر آب، روال سیگنال راهنما، کشف مسیر، و رفتار گره مخرب، را توضیح می‌دهیم.

۲-۱- مدل لینک ارتباطی

ما دو مدل برای لینک‌های ارتباطی بین هر جفت گره در نظر می‌گیریم.

اولی، سناریو ۱، یک مدل پایه مبتنی بر فاصله است. در این حالت، زمانی که فاصله بین آنها کمتر از حداکثر محدوده ارتباطی باشد، دو گره توسط یک پیوند ارتباطی معتبر به هم متصل می‌شوند. حداکثر محدوده ارتباطی یک پارامتر قابل تنظیم شبیه‌ساز است.

در حالت دوم، سناریو ۲، ما از یک مدل احتمالی از بهره کانال بین هر جفت گره استفاده می‌کنیم. به‌طور خاص، ما سود کانال در مقیاس بزرگ را با یک مؤلفه قطعی و یک مؤلفه تصادفی مدل می‌کنیم. جزء تصادفی با استفاده از فرآیند AR-۱ به‌عنوان یک فرآیند تصادفی عادی گزارشی مدل‌سازی می‌شود [۱۹]. سود کانال توسط فرمول زیر محاسبه می‌شود.

$$g(d.t) = \bar{g}(d) + g_{AR}(t)(dB) \quad (۱)$$

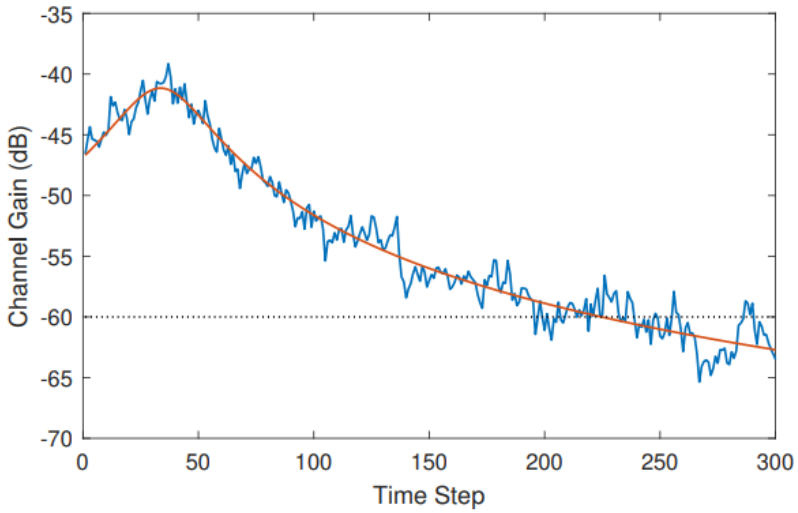
در اینجا، $\bar{g}(d)$ میانگین بهره لینک و $g_{AR}(t)$ فرایند تصادفی AR-1 (به دسی‌بل) است. میانگین بهره لینک به‌صورت قطعی براساس فاصله d بین گره‌ها و به شکل زیر محاسبه می‌شود:

$$\bar{g}(d) = g_0 - k_0 10\log 10 \quad (۲)$$

در اینجا، ثابت‌های g_0 و k_0 پارامترهای اتلاف انتشار بزرگ‌مقیاس هستند. مؤلفه بهره تصادفی به‌صورت فرایند تصادفی AR-1 و به شکل زیر مدل‌سازی می‌شود.

$$g_{AR}(t+\Delta) = \rho g_{AR}(t) + (1-\rho)\omega(t) \quad (۳)$$

در اینجا، Δ گام زمانی و ρ ضریب همبستگی در هر گام است که به اختلاف داپلر[1] f_0 براساس $\rho = (-2\pi f_0 \Delta)$ بستگی دارد، و $\omega(t)$ فرایند تصادفی گاوسی با میانگین صفر است. ما بر مقادیر تجربی ثابت‌های g_0 و k_0 تکیه می‌کنیم و f_0 از اندازه‌گیری‌های ثبت‌شده در آزمایش طوفان اقیانوس آرام (PS)[2] [۱۸] محاسبه می‌شود. یکی از مثال‌های عینی از بهره کانال در آزمایش PS در شکل ۱ نمایش داده شده است.



شکل ۱: خروجی شبیه‌سازی شده از بهره کانال برای سناریو ۲ (g۰ = -۳۹، K۰ = ۱.۸۹، f۰ = ۴۱۵، و Δ = ۳۰ ثانیه)

Figure 1 . Simulated output of channel gain for scenario 2 (g0 = -39, K0 = 1.89, f0 = 415, and Δ = 30 seconds)

۲-۲- مدل انرژی

مفروضات مربوط به اتلاف انرژی در حالت‌های ارسال و دریافت، مزیت‌های پروتکل‌های مختلف را تغییر خواهد داد. ما فرض می‌کنیم که اتلاف انرژی در اثر انتقال کانال وجود دارد. مدل انرژی برای گره‌های حسگر، هر گره دارای همان انرژی اولیه موجود برای ارسال یک پیام k-bit در فاصله R است، انرژی مصرف شده برای ارسال و دریافت پیام‌ها توسط رابطه زیر محاسبه می‌شود:

$$E_T X = E_{eleck} + E_{amp} k d^2 \quad (۴)$$

$$E_{RX} = E_{eleck} k \quad (۵)$$

انرژی مورد نیاز برای انتقال پیام E_{TX} است.

انرژی مورد نیاز برای دریافت پیام E_{RX} است.

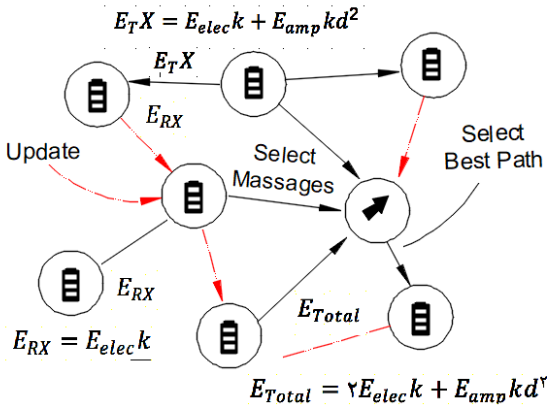
انرژی مورد نیاز برای مدار فرستنده گیرنده برای مقابله با یک بیت داده E_{elec} است.

انرژی مورد نیاز برای پردازش یک بیت داده به تقویت‌کننده فرستنده E_{amp} است.

d شعاع ارتباطی برابر گره‌ها است.

کانال رادیویی متقارن است به‌طوری‌که انرژی مورد نیاز برای انتقال یک پیام از گره A به گره B با انرژی مورد نیاز برای انتقال یک پیام از گره B به گره A برابر است.

[1] Doppler spread
[2] Pacific Storm



شکل ۲: مدل پیشنهادی انرژی
Figure 2. Energy Suggested Model

در این مدل انرژی، هر گره حسگر دارای منبع انرژی اولیه فرض می‌شود و هدف بهینه‌سازی عملکرد شبکه و در عین حال به حداقل رساندن مصرف انرژی برای افزایش طول عمر شبکه است. برای این مقادیر پارامتر، ارسال و دریافت پیام یک عملیات کم‌هزینه نیست. بنابراین پروتکل باید فواصل ارسال و همچنین تعداد عملیات ارسال و دریافت را برای هر پیام کاهش دهد. مسائلی که به آنها پرداخته می‌شود، محاسبه انرژی‌های ساخته شده با استفاده از مدل با توجه به فاصله بین گره‌ها و حذف گره‌های مرده از جدول همسایه است. در حین انتقال داده‌های جدید، اگر یک گره اطلاعاتی برای انتقال داشته باشد، گره در جدول مسیریابی خود بهترین همسایه را که نزدیک به سینک است بر اساس تعداد پرش و انرژی گره بررسی می‌کند. این کار با تعداد پرش شروع می‌شود و سپس انرژی بسته به اینکه گره بیش از یک گره در جدول همسایه خود با تعداد پرش یکسان داشته باشد و سپس انرژی باقیمانده را داشته باشد، انجام می‌شود. هنگامی‌که انرژی باقیمانده گره به زیر آستانه معینی می‌رسد، گره یک پیام به‌روزرسانی انرژی را به همسایگان خود ارسال می‌کند و آنها را از وضعیت انرژی خود مطلع می‌کند .هنگامی‌که گره‌ها پیام انرژی را دریافت می‌کنند، گره‌ای را که پیام را ارسال کرده است از جدول همسایه خود حذف می‌کنند.

۳- فازهای روش پیشنهادی (شبیه‌سازی)

برای شبیه‌سازی از نرم افزار NS2 استفاده شده است. NS2 به‌طور ساده یک ابزار شبیه‌سازی متن باز، مبتنی بر رویداد، وابسته به زمان و رخداد -گسسته است که در مطالعه طبیعت دینامیکی شبکه‌های ارتباطی مورد استفاده قرار می‌گیرد و از کتابخانه غنی از پروتکل‌های شبکه و آبجکت‌ها بهره می‌برد.

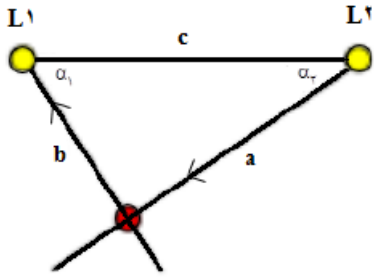
نوع ارسال بسته P2P بوده و ارسال و دریافت بسته به‌صورت همزمان نیز صورت می‌پذیرد. شبیه‌سازی از چهار مرحله تشکیل شده است: مرحله اول کشف همسایه امن تحت حملات کرم چاله در شبکه‌های حسگر زیر آب است. مرحله دوم فرآیند کشف مسیر اولیه است که گره ارسال قابل اعتماد بعدی را به گره سینک انتخاب می‌کند، بنابراین احتمال تحویل بسته به سمت سینک را افزایش می‌دهد و در نتیجه قابلیت اطمینان را افزایش می‌دهد. مرحله سوم فرآیند تشخیص حمله در حین توزیع داده است. این مرحله بر اساس اطلاعات وضعیت گره‌ها برای شناسایی حمله Sybil است. مرحله چهارم، کشف مسیر امن جایگزین برای شناسایی گره‌های مخرب است.

در شبکه‌های حسگر زیر آب که گره‌های حسگر ثابت و متحرک هستند، کشف همسایه یک نیاز اساسی است. در فرآیند کشف همسایه، هر گره همسایگان خود را کشف می‌کند و لیستی از همسایگان خود را ارائه می‌دهد. هنگامی‌که یک گره می‌خواهد گره‌های همسایه خود را پیدا کند، پیام درخواستی را در محدوده ارتباطی خود پخش می‌کند. هر پیام درخواستی شامل یک کلید خصوصی/کلید عمومی امضای دیجیتال، شماره شناسایی منحصر به فرد است. شماره شناسایی که در پیام درخواست تعبیه شده است پس از دریافت توسط گره مقصد در جدول ثبت می‌شود. با دریافت بسته‌های پخش، ابتدا گره گیرنده شماره شناسایی خود را در جدول جستجو می‌کند و در صورت وجود بسته دریافتی را حذف می‌کند. به‌این‌ترتیب از حملات مکرر جلوگیری می‌شود. اگر گره‌هایی که در محدوده ارتباطی گره بسته درخواست ارسال هستند موفق به تائید کلید عمومی/خصوصی در بسته

دریافت کننده شوند، تصمیم می‌گیرند گره‌های بسته‌ای را که دارای کلید عمومی/خصوصی معتبر هستند ارسال کنند و بسته پاسخ را به گره درخواست ارسال کنند. پیام پاسخ حاوی جهت سیگنال صوتی دریافتی و خود کلید عمومی و کلید خصوصی امضای دیجیتال است. برای یافتن جهت سیگنال صوتی، هر گره مجهز به آرایه هیدروفون است که می‌تواند جهت سیگنال صوتی ارسالی را تخمین بزند.

در صورت وجود تحرک گره در شبکه آکوستیک زیر آب، باید اثر آن بر روی گره‌ها بررسی شود. برای مدیریت گره‌های سیار، اثرات منفی این حرکات باید بر عملکرد پروتکل مسیریابی به حداقل برسد. مهم‌ترین روش کنترل تحرک، پیش‌بینی گره‌های متحرک است. در یک سناریوی واقعی، حرکات افقی در محدوده ۲-۳ متر بر ثانیه امکان‌پذیر نیست و فقط نوسانات جزئی وجود خواهد داشت، درحالی‌که به‌صورت عمودی، گره به‌طور مداوم با سرعت ۲-۳ متر بر ثانیه با جریان آب حرکت می‌کند. به‌عنوان مثال، گره x یک پیام درخواست ارسال می‌کند، گره y پیام را در L1 دریافت می‌کند، گره y بسته پاسخ را به گره x در L2 می‌فرستد. اکنون جهت سیگنال ارسالی از y به x تغییر کرده است. در چنین شرایطی، گره x پس از تائید امضای دیجیتال گره y، ابتدا بررسی می‌کند که معادله برقرار باشد.

با توجه به حرکت گره y از محل L1 به L2، یک مثلث بین گره x و گره y با رسم سیگنال‌های ارسالی و دریافتی همانطور که در شکل زیر نشان داده شده است ایجاد شد. a فاصله گره x تا گره y در L1 است. b فاصله گره x تا گره y در L2 است.



شکل ۳:تأثیر حرکت گره در فرآیند کشف همسایه (پیشنهادی)
Figure 3. The effect of node movement on neighbor discovery process

مسافت طی شده از محل گره قبلی تا مکان جدید C است.

$$\frac{a}{\sin a_1} = \frac{b}{\sin a_2} = 2R \quad (۶)$$

$$a^2 = b^2 + c^2 - 2bc \times \cos a_1 \quad (۷)$$

$$b^2 = a^2 + c^2 - 2ac \times \cos a_2 \quad (۸)$$

با به دست آوردن مقدار مساوی از معادله فوق می‌توان شعاع محیط مثلث را به دست آورد. اگر شعاع محیط مثلث بزرگتر از محدوده ارتباطی گره باشد، گره y از محدوده ارتباطی گره x خارج است و گره x گره y را به‌عنوان همسایه خود نمی‌پذیرد.

در صورت برقراری این رابطه، گره x مکان قبلی گره y را به‌روز می‌کند. جهت سیگنال صوتی از y تا x محاسبه می‌شود. گره y با دریافت بسته پاسخ از گره x، امضا را تائید می‌کند و سپس جهت دریافتی جدید سیگنال صوتی را از گره x محاسبه می‌کند و آن را در رابطه ۹ قرار می‌دهد.

$$\left|a_{yx} + a_{xy}\right| - \pi \le \partial \quad (۹)$$

δ خطاهای از پیش تعیین شده است. αxy زوایای جهت xy است. اگر ۱۸۰ درجه از مجموع سیگنال صوتی x به y کم شود و y به x کمتر از خطاهای از پیش تعیین شده باشد، گره y می‌تواند گره x را به‌عنوان همسایه واقعی بپذیرد و در لیست همسایگان خود قرار دهد و سپس بسته پاسخ را به گره x ارسال کند.

در بسته پاسخ، کلید عمومی گره x و y برای سیگنال صوتی قرار داده شده است. با دریافت بسته پاسخ توسط گره y، امضا تائید می‌شود. اگر امضا معتبر باشد، گره x تصمیم می‌گیرد که گره y دارای کلید عمومی/خصوصی معتبر باشد. سپس جهت سیگنال

صوتی y به x را محاسبه کرده و در رابطه قرار می‌دهد. در صورت برقراری رابطه فوق، گره x می‌تواند گره y را به‌عنوان یک همسایه واقعی بپذیرد و آن را در لیست همسایه خود قرار دهد.
پس از اینکه هر گره لیستی از همسایگان را ارائه کرد، فاصله فیزیکی واقعی بین دو گره محاسبه می‌شود تا روابط گره‌های همسایگی بررسی شود و کرم‌چاله شناسایی شود. اگر فاصله اندازه‌گیری شده بیشتر از محدوده گره‌های ارتباطی باشد، فرض می‌شود که گره‌ها از طریق کرم چاله به هم متصل شده‌اند. در SRAU، هر گره حسگر فاصله خود را با همسایگان محاسبه می‌کند. محاسبه گره مقصد با استفاده از انرژی سیگنال صوتی ارسالی توسط گره مبدأ و دریافت آن توسط گره مقصد انجام می‌شود. انرژی دریافتی به‌صورت معادله ۱۰ محاسبه می‌شود:

$$E_{recieve} = \frac{E_{send}}{4\,\pi R^2}\,\lambda \qquad (۱۰)$$

انرژی ارسال شده E_{send} است. R محدوده ارتباطی است و λ طول موج است. ما یک منطق فازی را در نظر می‌گیریم، بنابراین تابع ورودی فاصله و انرژی مصرفی گره است، تابع خروجی اعتبار سنجی گره است که جدول ۲ نشان‌دهنده آن هست.
در واقع هر گره همسایه‌هایی را که اعتبار پایینی دارند به‌عنوان حملات کرم چاله در نظر گرفته و از لیست همسایگان خود حذف می‌کند.

جدول ۱: پارامترهای شبیه‌سازی
Table 1. Simulation parameters

Parameter	Parameter Value
Network Dimensions	2000*2000
Number of nodes	40,50,70,100
Interface Queue Capacity	50 Packet
Interface Queue Type	DropTail
The initial energy of each node	50 Jol
Simulation Time	1500
Standard Type MAC	802.11

اطلاعات ورودی از مطالعات قبلی[۵] [۱۹] در جدول پارامترهای شبیه‌سازی ، محاسبه و بعد از استخراج خروجی‌ها، اطلاعات ورودی شامل تعداد گره و فضای شبیه‌سازی و زمان شبیه‌سازی، اطلاعات کیفیت سرویس از این ورودی‌ها استخراج می‌شود.
ازآنجاکه مسیریابی یک مسئله اساسی در شبکه است، الگوریتم مسیریابی باید گره ارسال بعدی را انتخاب کند که احتمال تحویل بسته را به سمت سینک افزایش می‌دهد و در نتیجه شاخص متریک قابلیت اطمینان که به لینک‌های شبکه مرتبط بوده و به‌صورت داینامیک محاسبه می‌شود را افزایش می‌دهد. هر گره یک جدول مسیریابی محلی دارد . هنگامی‌که هر گره حسگر یک بسته را ارسال می‌کند، یک شناسه منحصر به فرد به آن متصل می‌شود.
در فرآیند مسیریابی، ابتدا یک شاخص اتصال توسط گره سینک به هر گره اختصاص داده می‌شود. Sink node ابتدا یک بسته hello را پخش می‌کند تا ایندکس را به هر گره اختصاص دهد. به‌محض دریافت بسته‌های hello توسط گره‌های حسگر، یک شاخص اتصال و تعداد پرش از گره سینک تخصیص داده می‌شود. سپس با دریافت بسته‌های hello توسط گره‌ها، شاخص اتصال و تعداد پرش آنها به سمت سینک بازپخش می‌شود که ممکن است این بسته را توسط گره‌های همسایه خود نیز دریافت کنند. هنگامی‌که یک گره بسته‌ای را از گره همسایه دریافت می‌کند، ابتدا تعداد پرش گره‌های همسایه بررسی می‌شود. اگر تعداد پرش گره‌های همسایه کمتر یا مساوی از تعداد پرش آن باشد، شاخص اتصال آن یک واحد افزایش می‌یابد و شناسه، شاخص اتصال و تعداد پرش گره‌های همسایه را حفظ می‌کند. اکنون، هر گره یک بسته برای ارسال به گره سینک دارد و باید گره ارسال بعدی را انتخاب کند. ابتدا، آن گره لیست همسایه را بر اساس بالاترین شاخص اتصال در بین همسایگان خود که لیست همسایه آنها در مرحله قبل به دست آمده است مرتب می‌کند. اکنون، گره ارسال بعدی انتخاب شده بر اساس بالاترین شاخص اتصال، تعداد

پرش کوچکتر و انرژی باقی‌مانده بیشتر است. سپس، گره بسته درخواستی را به گره انتخابی بعدی ارسال می‌کند که حاوی موقعیت گره و مقصد نهایی است و منتظر بسته تائید می‌شود و بسته خود را با بسته تائید دریافتی ارسال می‌کند. عمر مسیریابی به زمان مهلت یا تائید دریافتی مربوطه بستگی دارد به‌طوری‌که هر گره با یک تایمر تنظیم می‌شود و منتظر بسته تائید است. اگر زمان منقضی شود یا بسته خراب دریافت کند، بخش دوباره ارسال می‌شود.

جدول ۲: تائید اعتبار
Table 2. Validation

Rate of validation	Energy consumption	distance
A lot	low	low
A lot	A lot	low
low	low	A lot
low	A lot	A lot

از شاخص اتصال متریک استفاده گردیده است که :

- متریک یک مقدار اختصاص داده شده به هر مسیر است.
- این مقدار تعیین می‌کند که مسیرهای بهینه برای ارسال بسته‌ها انتخاب می‌شوند.
- معمولاً متریک تعداد پرش‌ها (HOP) یا تعداد روترهایی است که باید برای رسیدن به شبکه مقصد از آن‌ها عبور کرد.
- اگر چندین مسیر وجود داشته باشد، معمولاً مسیری با کمترین متریک انتخاب می‌شود.
- مکانیزم جدول مسیریابی به این صورت است که وقتی بسته‌ای از مبدأ به مقصد ارسال می‌شود، روتر آن را باز می‌کند و اگر آدرس IP مقصد در جدول مسیریابی قرار داشت، بسته را به شبکه موردنظر هدایت می‌کند .در غیر این صورت، به بهترین مسیر برای رسیدن به مقصد هدایت می‌شود.
- این فرآیند باعث ارسال بهینه‌تر بسته‌ها در شبکه می‌شود.
- شماره شناسه:(Node ID)

این شماره به هر گره در شبکه اختصاص داده می‌شود که از آن برای شناسایی گره‌ها در جدول مسیریابی استفاده می‌شود.

- اتصال شاخص به سینک:(Connectivity to Sink)

این پارامتر نشان‌دهنده اتصال گره به سینک (گره مقصد) است. می‌تواند مقدار بولی (بله/خیر) داشته باشد.

- تعداد گام مانده به سینک:(Hops to Sink)

تعداد گام‌های باقی‌مانده تا رسیدن به سینک را نشان می‌دهد. این پارامتر می‌تواند عدد صحیح نامنفی باشد.

- انرژی:(Energy)

میزان انرژی باقی‌مانده در گره را نشان می‌دهد. معمولاً از واحدهای میلی‌ژول یا ژول استفاده می‌شود.

- طول عمر گره:(Node Lifetime)

مدت زمانی که گره قابل‌استفاده است را نشان می‌دهد. می‌تواند به واحدهای زمانی مانند ثانیه، دقیقه یا ساعت باشد.

- این پارامترها هنگام اجرا مقداردهی می‌شود:

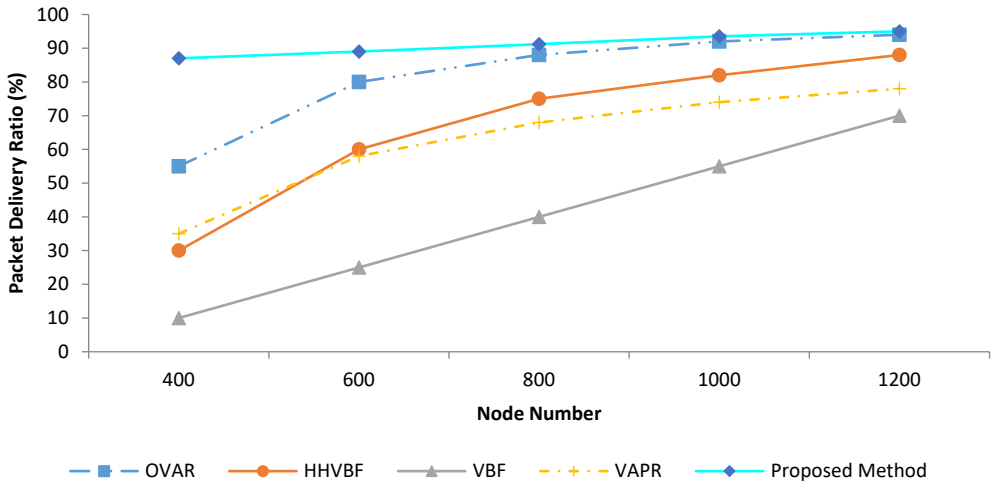
گره همسایه با دریافت هر قسمت، داده‌ها را بررسی می‌کند و قسمت در نظر گرفته شده را مجدداً به مقصد ارسال می‌کند. در طول انتشار داده‌ها، یک گره اغلب با گره‌های دیگر ارتباط برقرار می‌کند و هر گره هنگام ارسال و دریافت بسته‌ها انرژی مصرف می‌کند. سپس، انرژی گره باقی‌مانده کاهش می‌یابد. برخی ارزیابی‌ها برای شناسایی حمله Sybil در حین انتشار داده‌ها انجام می‌شود. حملات Sybil یک تهدید جدی در شبکه‌های حسگر زیر دریایی‌ها هستند. در چنین حملاتی، یک گره مخرب چندین هویت جعلی برای خود ایجاد می‌کند و گره‌های شبکه را گمراه می‌کند و انرژی گره‌های حسگر باقیمانده را کاهش می‌دهد. بنابراین طول عمر شبکه کاهش می‌یابد. برای یافتن گره مشکوک، داده‌های هر گره بررسی می‌شود.

در یک دوره زمانی مشخص، گره x شروع به ارسال بسته ها به گره y می‌کند. بسته‌های ارسال شده با Psend نمایش داده می‌شود.

گره y بسته ها را از گره x دریافت می‌کند. بسته‌های ارسال شده با Precieve نمایش داده می‌شود.
گره x پس از دریافت پاسخ از گره y، زمان اتصال را ثبت می‌کند. زمان پایان اتصال با Tend نمایش داده می‌شود.

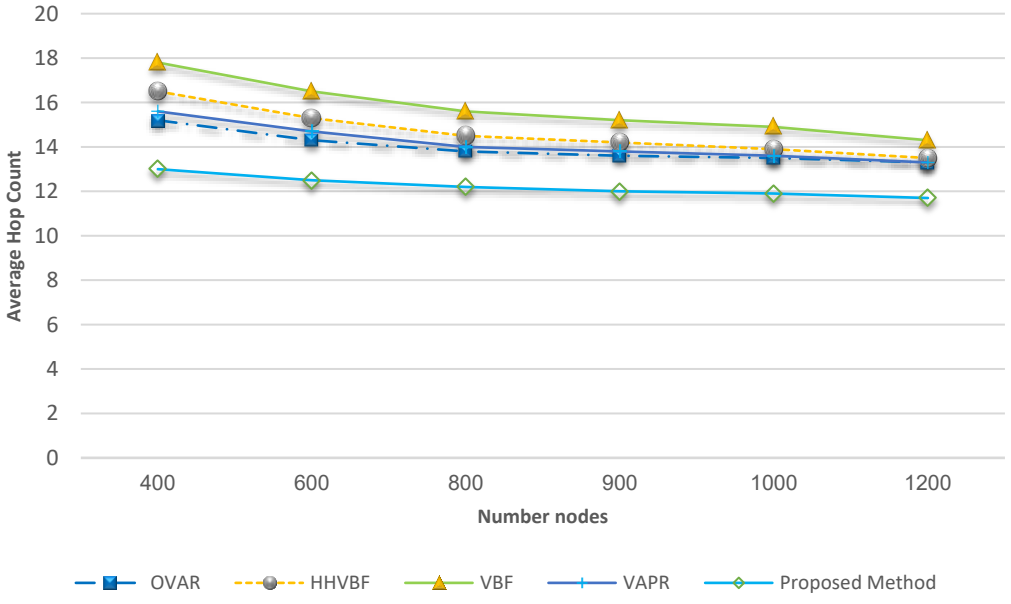
۴- نتایج شبیه‌سازی

برای هر خروجی ثبت شده در نمودار ۴۰ بار اجرا گرفته شده و میانگین آن به دست آمده است.



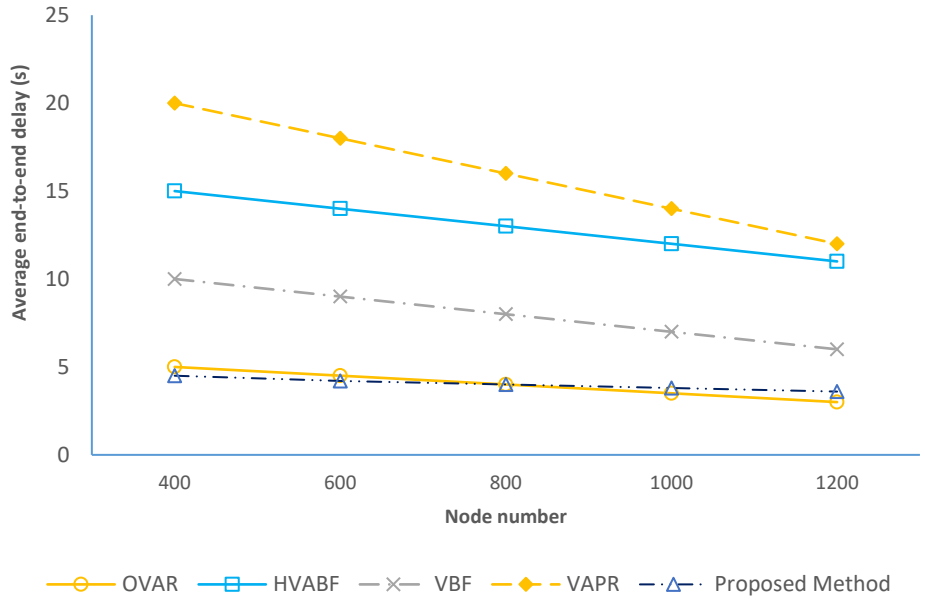
شکل ۴: مقایسه نمودار نرخ تحویل بسته شبکه نسبت به تعداد گره با مطالعات قبلی [۶]
Figure 4. Comparison of the graph of network packet delivery rate relative to the number of nodes with previous studies[6]

نرخ تحویل بسته، نسبت تعداد پیام‌های داده‌ای است که با موفقیت منتقل شده‌اند و توسط مقصد دریافت شده‌اند. متریک اول ما اندازه‌گیری نسبت موفقیت میانگین برای تعداد گره‌های مختلف متحرک می‌باشد. تمام پارامترهای شبیه‌سازی دیگر، مانند سناریوی شبیه‌سازی پیش فرض، مشخصات دهی شده‌اند. همانطور که در شکل ۴ نشان داده شده است میانگین نسبت موفقیت افزایش می‌یابد. نتیجه‌ی ما نشان می‌دهد که طرح پیشنهادی به نرخ موفقیت بهتری نسبت به طرح پایه[۶] دست یافته و با هزینه‌ی انرژی تحرک کمتری نسبت به روش پایه ، نسبت موفقیت قابل‌مقایسه‌ای را به دست آورده است.



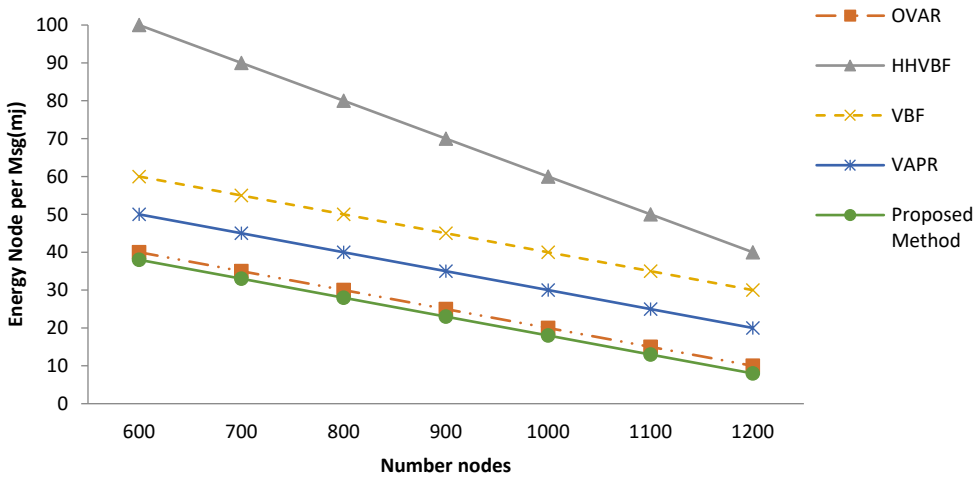
شکل ۵ : مقایسه نمودار توان عملیاتی شبکه نسبت به تعداد گره و تعداد پرش‌ها با مطالعات قبلی [۶]
Figure 5. Comparison of the graph of network throughput relative to the number of nodes and the number of nodes with previous studies[6]

ارتباط بین توان عملیاتی و تعداد پرش نود به این صورت است که با افزایش تعداد پرش‌ها، احتمال از دست رفتن داده‌ها و تأخیر در انتقال افزایش می‌یابد، که می‌تواند منجر به کاهش توان عملیاتی شود. به عبارت دیگر، هرچه تعداد پرش‌ها بیشتر باشد، احتمال بروز مشکلاتی مانند ازدحام و تأخیر بیشتر می‌شود، که این عوامل می‌توانند توان عملیاتی شبکه را کاهش دهد. شکل ۵ نشان می‌دهد که روش پیشنهادی نسبت به روش پایه دیگر [۶] دارای بیش‌ترین توان عملیاتی شبکه بوده است.



شکل ۶: مقایسه نمودار تاخیر انتها به انتها شبکه نسبت به تعداد گره با مطالعات قبلی [۶]

Figure 6. Comparison of the end-to-end delay chart of the network relative to the number of nodes with previous studies[6]



شکل ۷: مقایسه نمودار مصرف انرژی کل شبکه نسبت به تعداد گره با انرژی مصرفی هر گره با مطالعات قبلی [۶]

Figure 7. Comparing the graph of energy consumption of the entire network with the number of nodes and the energy consumption of each node with previous studies[6]

در نمودارشکل ۶، میانگین تأخیر انتها به انتها را به‌عنوان نموداری از تعداد گره‌های متحرک با سرعت متفاوت بررسی و با مطالعات قبل [۶] مقایسه نمودیم. این معیار میانگین زمان تأخیر گرفته شده از لحظه ایجاد بسته ها در مبدأ تا دریافت توسط سینک را برای همه بسته ها با موفقیت محاسبه می‌کند. تاخیر انتشار، تاخیر انتقال و زمان نگهداری بسته ها را برای محاسبه تاخیر انتها به انتها در نظر می‌گیریم. متوسط تأخیر انتها به انتها برای همه پروتکل‌ها با افزایش تعداد گره‌ها کاهش می‌یابد و به‌صورت کلی تاخیر نسبت به مطالعات قبلی کاهش یافته است. در این تحلیل، میانگین تاخیر به ازای هر گره در شبکه تأثیرگذار

است. اگر چه تغییر مسیر دینامیکی از یک فرستنده، مسیر مختلفی را به مقصدهای متحرک که با سرعت‌های مختلف حرکت می‌کنند، می‌فرستد.

در نمودار شکل ۷، میانگین مصرف انرژی بررسی شده است. تراکم گره، تأثیر کمی بر انرژی به ازای هر گره در روش پیشنهادی دارد، اگر چه همسایه‌های بیشتری، داده‌های یک فرستنده‌ی با تراکم بالا را استراق سمع می‌کنند. به این دلیل که شانس بیشتری وجود دارد که مسیرهای بهتر از نظر هزینه‌ی انرژی را می‌توان در یک شبکه‌ی با تراکم بالاتر یافت. روش پیشنهادی در شکل ۷، در مصرف انرژی کل، عملکرد بهتری رانسبت به مطالعات قبلی[۶] نشان می‌دهد. علت اینکه به مصرف انرژی کمتری دست یافته است این است که منبع می‌تواند رویداد انتقال را از مبدأها به مقصدهای با مسیرهای مناسب هدایت کند.

۵- نتیجه‌گیری

در مقاله حاضر، شبیه‌سازی کشف همسایه و مسیر در شبکه ارتباطات صوتی زیر آب را ارائه کردیم.کاربردی از اینترنت اشیای زیر آب را در نظر گرفتیم که در آن توانمندی‌های پردازشی برای رمزگذاری/هویت‌سنجی و توانمندی‌های ارتباطی برای تصدیق یا بازانتقال ممکن است محدود باشند. انتقال سیگنال‌های راهنما و تشکیل جداول همسایه را تحت دو مدل لینک ارتباطی شبیه‌سازی کردیم. رفتار گره مخرب بر کشف مبتنی بر سیگنال راهنما را اندازه‌گیری کردیم. نتایج ما نشان می‌دهد که گره مخرب می‌تواند با موفقیت بر عملکرد شبکه تأثیر بگذارد. نتایج مطابق جدول زیر با مطالعات قبلی مقایسه شده است و بیانگر موفقیت در توان عملیاتی، تاخیر انتها به انتها، نرخ تحویل بسته و تا حدودی مصرف انرژی است.

جدول۳: مقایسه نرخ تحویل بسته شبکه به تعداد گره در مطالعه جاری و مطالعات قبلی

Table 3. Comparing the network packet delivery rate to the number of nodes in the current study and previous studies

Porotocol	Ref	year	Packet delivery rate	Energy consumption of each node	End-to-end delay of the network(s)
VBF	[27]	2006	66	14	10/5
HHVBF	[26]	2007	90	20	-
FVBF	[25]	2018	72	7/5	6/4
EAVARP	[28]	2018	92	6	2/3
RDBF	[29]	2013	80	15	6
GEDAR	[30]	2016	82	15	5
SEECR	[31]	2020		33	23
GRMC-SM	[32]	2018	90	14	5/3
DMR	[33]	2019	35	7	-
SUGGESTED METHOD	-	-	94	8	3.2

مراجع

[1] B. D. Deebak and F. Al-Turjman, "A hybrid secure routing and monitoring mechanism in IoT-based wireless sensor networks," *Ad Hoc Networks*, vol. 97, p. 102022, 2020. doi: 10.1016/j.adhoc.2019.102022

[2] S. Nepali, "The Secure and Energy Efficient Data Routing in the IoT-based Network," 2020. doi: 10.1016/J.ADHOC.2019.102022

[3] J. Luo, Y. Chen, M. Wu, and Y. Yang, "A survey of routing protocols for underwater wireless sensor networks," *IEEE Commun. Surveys & Tutorials*, vol. 23, no. 1, pp. 137-160, 2021. doi: 10.1109/COMST.2020.3048190

[4] S. M. Ghoreyshi, A. Shahrabi, and T. Boutaleb, "Void-handling techniques for routing protocols in underwater sensor networks: Survey and challenges," *IEEE Commun. Surveys & Tutorials*, vol. 19, no. 2, pp. 800-827, 2017. doi: 10.1109/COMST.2017.2657881.

[5] D. J. Jakubisin, C. McPeak, J. Sloop, and B. Davis, "Securing route discovery for the underwater Internet of Things," *OCEANS 2022, Hampton Roads*, pp. 1-10, Oct. 2022. doi: 10.1109/OCEANS47191.2022.9977183.

[6] S. M. Ghoreyshi, A. Shahrabi, and T. Boutaleb, "A novel cooperative opportunistic routing scheme for underwater sensor networks," *Sensors*, vol. 16, no. 3, p. 297, 2016. doi**:** 10.3390/s16030297

[7] H. Kaushal and G. Kaddoum, "Underwater optical wireless communication," *IEEE Access*, vol. 4, pp. 1518-1547, 2016. doi: 10.1109/ACCESS.2016.2552538

[8] G. Yang, L. Dai, and Z. Wei, "Challenges, threats, security issues, and new trends of underwater wireless sensor networks," *Sensors*, vol. 18, no. 11, p. 3907, 2018. doi: 10.3390/s18113907

[9] G. Han, J. Jiang, N. Sun, and L. Shu, "Secure communication for underwater acoustic sensor networks," *IEEE Commun. Mag.*, vol. 53, no. 8, pp. 54-60, 2015. doi: 10.1109/MCOM.2015.7180508

[10] M. C. Domingo, "Securing underwater wireless communication networks," *IEEE Wireless Commun.*, vol. 18, no. 1, pp. 22-28, 2011. doi: 10.1109/MWC.2011.5714022

[11] A. G. Yisa, T. Dargahi, S. Belguith, and M. Hammoudeh, "Security challenges of internet of underwater things: A systematic literature review," *Trans. Emerg. Telecommun. Technol.*, vol. 32, no. 3, p. e4203, 2021. doi: 10.1002/ETT.4203

[12] H. H. Rizvi, R. N. Enam, S. A. Khan, and J. Akram, "A survey on internet of underwater things: Perspective on protocol design for routing," *2020 Global Conf. on Wireless and Optical Technologies (GCWOT)*, pp. 1-8, Oct. 2020. doi: 10.1109/GCWOT49901.2020.9391628

[13] N. Javaid, S. Cheema, M. Akbar, N. Alrajeh, M. S. Alabed, and N. Guizani, "Balanced energy consumption-based adaptive routing for IoT enabling underwater WSNs," *IEEE Access*, vol. 5, pp. 10040-10051, 2017. doi: 10.1109/ACCESS.2017.2706741.

[14] S. M. Ghoreyshi, A. Shahrabi, and T. Boutaleb, "An inherently void avoidance routing protocol for underwater sensor networks," *2015 International Symposium on Wireless Communication Systems (ISWCS)*, pp. 361-365, Aug. 2015. doi: 10.1109/ISWCS.2015.7454364

[15] Y. Noh, U. Lee, P. Wang, B. S. C. Choi, and M. Gerla, "VAPR: Void-aware pressure routing for underwater sensor networks," *IEEE Trans. Mobile Comput.*, vol. 12, no. 5, pp. 895-908, 2012. doi: 10.1109/TMC.2012.53

[16] R. Zhang and Y. Zhang, "Wormhole-resilient secure neighbor discovery in underwater acoustic networks," *2010 Proceedings IEEE INFOCOM*, pp. 1-9, Mar. 2010. doi: 10.1109/INFCOM.2010.5462093

[17] F. Jan, N. Min-Allah, and D. Düştegör, "IoT-based smart water quality monitoring: Recent techniques, trends, and challenges for domestic applications," *Water*, vol. 13, no. 13, p. 1729, 2021. doi: 10.3390/w13131729

[18] N. Adam, M. Ali, F. Naeem, A. S. Ghazy, and G. Kaddoum, "A comprehensive survey of security schemes and privacy-preserving techniques for the internet of underwater things," *Communications*, vol. 1, no. 2, 2024. doi: 10.36227/TECHRXIV.171502698.82158311

[19] V. G. Menon and P. J. Prathap, "Comparative analysis of opportunistic routing protocols for underwater acoustic sensor networks," *2016 International Conference on Emerging Technological Trends (ICETT)*, pp. 1-5, Oct. 2016. doi: 10.1109/ICETT.2016.7873733

[20] M. T. Kheirabadi and M. M. Mohamad, "Greedy routing in underwater acoustic sensor networks: a survey," *Int. J. Distrib. Sens. Netw.*, vol. 9, no. 7, p. 701834, 2013. doi: 10.1155/2013/701834

[21] M. Chaudhary, N. Goyal, and A. Mushtaq, "Internet of underwater things: challenges, routing protocols, and ML algorithms," *Machine Learning Paradigm for Internet of Things Applications*, pp. 247-263, 2022. doi: 10.1002/9781119763499.CH13

[22] E. C. Liou, C. C. Kao, C. H. Chang, Y. S. Lin, and C. J. Huang, "Internet of underwater things: Challenges and routing protocols," *2018 IEEE International Conference on Applied System Invention (ICASI)*, pp. 1171-1174, Apr. 2018. doi: 10.1109/ICASI.2018.8394494

[23] S. M. Ghoreyshi, A. Shahrabi, and T. Boutaleb, "A stateless opportunistic routing protocol for underwater sensor networks," *Wireless Commun. Mob. Comput.*, vol. 2018, p. 8237351, 2018. doi: 10.1155/2018/8237351

[24] P. Qarabaqi and M. Stojanovic, "Modeling the large-scale transmission loss in underwater acoustic channels," *2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 445-452, Sep. 2011. doi: 10.1109/ALLERTON.2011.6120201.

[25] R. Bu, S. Wang, and H. Wang, "Fuzzy logic vector-based forwarding routing protocol for underwater acoustic sensor networks," *Trans. Emerg. Telecommun. Technol.*, vol. 29, no. 3, p. e3252, 2018. doi: 10.1002/ETT.3252.

[26] N. Nicolaou, A. See, P. Xie, J. H. Cui, and D. Maggiorini, "Improving the robustness of location-based routing for underwater sensor networks," *Oceans 2007-Europe*, pp. 1-6, Jun. 2007. doi: 10.1109/OCEANSE.2007.4302470.

[27] P. Xie, J. H. Cui, and L. Lao, "VBF: Vector-based forwarding protocol for underwater sensor networks," *NETWORKING 2006. Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications Systems: 5th International IFIP-TC6 Networking Conference*, Coimbra, Portugal, May 15-19, 2006, Proc., vol. 5, pp. 1216-1221. Springer Berlin Heidelberg. doi: 10.1007/11753810_111

[28] Z. Wang, G. Han, H. Qin, S. Zhang, and Y. Sui, "An energy-aware and void-avoidable routing protocol for underwater sensor networks," *IEEE Access*, vol. 6, pp. 7792-7801, 2018. doi: 10.1109/ACCESS.2018.2805804

[29] Z. L. Li, N. M. Yao, and Q. Gao, "Relative distance-based forwarding protocol for underwater wireless sensor networks," *Applied Mechanics and Materials*, vol. 437, pp. 655-658, 2013. doi: 10.4028/WWW.SCIENTIFIC.NET/AMM.437.655

[30] R. W. Coutinho, A. Boukerche, L. F. Vieira, and A. A. Loureiro, "Geographic and opportunistic routing for underwater sensor networks," *IEEE Trans. Comput.*, vol. 65, no. 2, pp. 548-561, 2015. doi: 10.1109/TC.2015.2423677

[31] K. Saeed, W. Khalil, S. Ahmed, I. Ahmad, and M. N. K. Khattak, "SEECR: Secure energy efficient and cooperative routing protocol for underwater wireless sensor networks," *IEEE Access*, vol. 8, pp. 107419-107433, 2020. doi: 10.1109/ACCESS.2020.3000863

[32] F. Ahmed, Z. Wadud, N. Javaid, N. Alrajeh, M. S. Alabed, and U. Qasim, "Mobile sinks assisted geographic and opportunistic routing-based interference avoidance for underwater wireless sensor networks," *Sensors*, vol. 18, no. 4, p. 1062, 2018. doi: 10.3390/s18041062

[33] U. Ullah, A. Khan, S. M. Altowaijri, I. Ali, A. U. Rahman, V. V. Kumar, ... and H. Mahmood, "Cooperative and delay minimization routing schemes for dense underwater wireless sensor networks," *Symmetry*, vol. 11, no. 2, p. 195, 2019. doi: 10.3390/SYM11020195.

JOURNAL OF SOUTHERN COMMUNICATION ENGINEERING
ISLAMIC AZAD UNIVERSITY BUSHEHR BRANCH

E-ISSN: 2980-9231
https://sanad.iau.ir/journal/jce

https://doi.org/10.30495/jce.2025.1993480.1337

Vol. 14/ No. 56/2025

Research Article

A New BTS Antenna for Simultaneous Operation in 900 GSM Frequency Bands and LTE

Maryam Mohammadifar[1] | Pejman Mohammadi[2*] |Yashar Zehforoosh[3]

[1]Department of Electrical Engineering, Urmia Branch, Islamic Azad University, Urmia Iran.
mohammadifar_eng@yahoo.com

[2]Microwave and antenna research center, Urmia Branch, Islamic Azad University, Urmia, Iran.
P.mohammadi@iaurmia.ac.ir

[3]Microwave and antenna research center, Urmia Branch, Islamic Azad University, Urmia, Iran.
y.zehforoosh@srbiau.ac.ir

Correspondence
Pejman Mohammadi, Associate Professor of the Microwave and Antenna Research Center, Urmia Branch, Islamic Azad University, Urmia, Iran.
P.mohammadi@iaurmia.ac.ir

Main Subjects:
Antenna Design

Paper History:
Received: 25 July 2024
Revised: 15 August 2024
Accepted: 3 September 2024

Abstract

This paper presents a novel dual-band, dual-polarized base station antenna design for mobile communication systems operating across GSM/DCS/PCS/UMTS and LTE frequency bands. The antenna achieves wide input impedance matching bandwidth through an innovative trident-shaped feeding technique. The design consists of two printed dipoles positioned perpendicular to each other and fed by stepped-microstrip lines. A low-profile cavity-backed structure serves as a metal reflector beneath the antenna, effectively converting the dipoles' bidirectional radiation patterns into unidirectional patterns while simultaneously increasing antenna gain. The proposed antenna demonstrates effective operation at 900, 1800, 1900, and 2300 MHz frequency bands, making it suitable for base station applications. Measurement results show peak gains of 11.47 dBi at port-1 and 10.40 dBi at port-2. The antenna element itself measures 168×168mm² and is mounted on a 222×222mm² cavity-backed structure with a depth of 42 mm. These features combine to create a compact yet high-performance solution for modern multi-standard cellular systems, with the dual-polarized configuration and broadband characteristics proving particularly advantageous for current communication infrastructure requirements.

Keywords: Base Station Antenna, Dual-Polarization, Printed Antenna, Trident-Shaped Feeding.

Highlights

- Simple antenna structure compared to previous designs.
- New dual-band antenna design with dual polarization for simultaneous operation in 900GSM and LTE frequency bands.
- Good and high gain compared to previous designs.

Citation: M. Mohammadifar, P. Mohammadi, and Y. Zehforoosh, "A New BTS Antenna for Simultaneous Operation in 900 GSM Frequency Bands and LTE," *Journal of Southern Communication Engineering*, vol. 14, no. 56, pp. 75–88, 2025, doi:10.30495/jce.2025.1993480.1337 [in Persian].

1. Introduction

Signal fading in mobile communications remains a critical challenge for antenna designers, primarily caused by wave scattering from surrounding objects like buildings, vehicles, and trees. This phenomenon necessitates the use of multiple antennas to enhance channel capacity and mitigate fading effects. However, implementing multiple antennas introduces mutual coupling issues, traditionally addressed through space diversity solutions that require larger physical dimensions. Recent advancements in wireless communication systems demand more compact designs with reduced installation space, lower costs, and support for multiple frequency bands, driving the need for dual-polarized, dual-band antennas.

Designing antennas with both dual-polarization and dual-band capabilities presents significant challenges. Conventional single-band dual-polarized antennas typically employ two orthogonal dipoles to achieve vertical and horizontal linear polarizations [1], but these cannot meet the comprehensive requirements of modern communication systems. Recent developments have focused on dual-band dual-polarized antennas for base station applications [2]-[3]. For instance, Deillon et al. [2] presented a design covering GSM, DCS, and UMTS bands using four microstrip-fed slots with corresponding cavities, achieving bandwidths of 5.6% (0.941-0.995 GHz) and 34.9% (1.710-2.434 GHz) for port-1, and 6.2% (0.932-0.992 GHz) and 19.8% (1.702-2.077 GHz) for port-2. Similarly, Ro and Huan [3] implemented a dual-band dual-polarized WLAN antenna using stacked microstrip structures. Additional innovations include novel designs for dual-polarized BTS antennas [4]-[7], small-cell BTS solutions [8], next-generation mobile systems [9]-[10], and microcell base stations [11].

This paper introduces a new base station antenna design featuring two printed orthogonal microstrip dipoles and a low-profile cavity-backed structure. The innovative trident-shaped microstrip feeding technique [12]-[13] enables measured impedance bandwidths of 30.17% (0.729-0.988 GHz) and 60.95% (1.289-2.419 GHz) for port-1, and 26.30% (0.743-0.968 GHz) and 32.11% (1.746-2.414 GHz) for port-2. Meeting contemporary BTS system requirements for compact, low-profile multifunctional antennas [14], the proposed design occupies a minimal volume of 222 × 222 × 42 mm³ while delivering dual-band operation and dual-polarization capabilities.

2. Innovation and contributions

The key innovations of this work include:

- A simplified antenna structure offering improved manufacturability compared to previous complex designs
- A novel dual-band, dual-polarized configuration enabling simultaneous operation in both 900MHz GSM and LTE frequency bands
- Enhanced gain performance surpassing existing antenna designs in this class

The proposed antenna maintains a compact size while improving key performance metrics.

3. Materials and Methods

The antenna structures were simulated using HFSS Microwave Studio software.

4. Results and Discussion

Figure I presents the simulated results of the proposed antenna obtained using ANSYS HFSS software. When both ports are driven, they demonstrate identical impedance bandwidths, covering 852-986 MHz and 1760-2463 MHz frequency ranges. The simulated isolation between ports exceeds 19dB.

The triple microstrip feeding technique significantly enhances the antenna's impedance bandwidth through several mechanisms. First, it increases the electrical length of the feeding structure, enabling multi-frequency resonance. Second, it introduces additional resonant modes that collectively expand the operational bandwidth. Third, the geometry incorporates both inductive and capacitive components that can be optimized for impedance matching across multiple frequencies.

This feeding configuration also modifies the current distribution on the radiating patch, promoting more uniform current flow over a broader frequency range. As shown in Figure II, the triple feed structure provides substantially wider bandwidth compared to a single feed, particularly in higher frequencies. This design approach effectively covers all essential BTS operating bands (800/900/1800/1900/2300 MHz) while maintaining excellent impedance matching characteristics.

5. Conclusion

This paper presents a novel base station antenna design featuring orthogonal dipoles with triple-feed series and a cavity reflector structure. The antenna achieves dual-polarization operation across two frequency bands while covering multiple mobile communication bands. Key design features include:

1. A triple-feed technique that significantly enhances impedance bandwidth
2. A cavity reflector structure that converts the radiation pattern to unidirectional
3. Stable radiation characteristics with less than 2dB gain variation at both ports

The antenna is fabricated on a 168×168mm² Rogers RO4003 substrate, complemented by a 222×222×42mm³ cavity reflector. Simulation results demonstrate:

- Port isolation exceeding 19dB
- Maximum gains of 11.40dBi (Port 1) and 10.47dBi (Port 2)
- Consistent performance across all operational bands

This compact design successfully addresses key requirements for modern base station antennas, including wide bandwidth, high isolation, and stable radiation patterns.

6. Acknowledgement

The authors gratefully acknowledge the Islamic Azad University of Urmia, Iran, for providing access to the PNA-X network analyzer system (N5242A) in their microwave laboratories for antenna testing purposes.

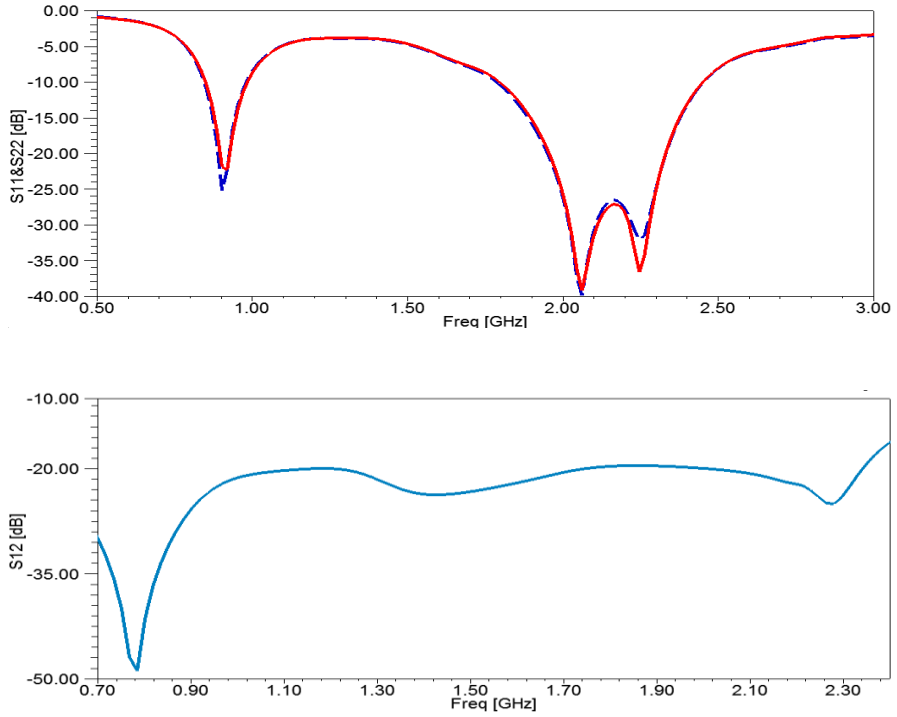


Figure 1. Simulated S parameters of the proposed antenna

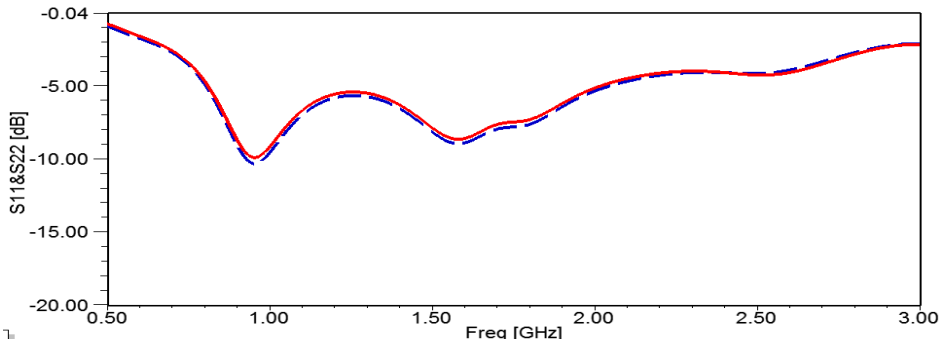


Figure II. Return loss of the proposed antenna with a single feed

References

[1] B. Lindmark and M. Nilsson, "On the available diversity gain from different dual-polarized antennas," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 2, pp. 287-294, 2001, doi: 10.1109/49.914506.

[2] S. Qu, J. Li, and Q. Xue, "Bowtie Dipole Antenna with Wide Beamwidth for Base Station Application," *IEEE Antennas and Wireless Propagation Letters*, vol. 6, pp. 293-295, 2007, doi: 10.1109/LAWP.2007.898543.

[3] W. X. An, H. Wong, K. L. Lau, S. F. Li, and Q. Xue, "Design of Broadband Dual-Band Dipole for Base Station Antenna," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 3, pp. 1592-1595, 2012, doi: 10.1109/TAP.2011.2180336.

[4] H. Zhai, Q. Gao, C. Liang, R. Yu, and S. Liu, "A Dual-Band High-Gain Base-Station Antenna for WLAN and WiMAX Applications," *IEEE Antennas and Wireless Propagation Letters*, vol. 13, pp. 876-879, 2014, doi: 10.1109/LAWP.2014.2321503.

[5] M. Rezvani and P. Mohammadi, "Microstrip antenna with aperture reflector and C-shaped dipoles for LTE and wireless communications," *Int J Electron Commun (AEÜ),* vol. 94, pp.12-18, 2018, doi: 10.1016/j.aeue.2018.06.041.

[6] Q. X. Chu, D. L. Wen, and Y. Luo, "A Broadband ± 45º Dual-Polarized Antenna With Y-Shaped Feeding Lines," *IEEE Trans. Antennas Propag.*, vol. 63, no. 2, pp. 483-490, 2015, doi: 10.1109/TAP.2014.2381238.

[7] Y. Gou, S. Yang, J. Li, and Z. Nie, "A compact dual-polarized printed dipole antenna with high isolation for wideband base station applications," *IEEE Trans. Antennas Propag*., vol. 62, no. 8, pp. 4392-4395, 2014, doi: 10.1109/TAP.2014.2327653.

[8] A. Elsherbini, J. Wu, and K. Sarabandi, "Dual polarized wideband directional coupled sectorial loop antennas for radar and mobile base-station applications," *IEEE Trans. Antennas Propag*., vol. 63, no. 4, pp. 1505-1513, 2015, doi: 10.1109/TAP.2015.2392773.

[9] V. Deillon and J-F. Zürcher, and A. K. Skrivervik, "A compact dual-band dual-polarized antenna element for GSM/DCS/UMTS base stations," *Microwave and Optical Technology Letters*., vol. 40, no. 1, pp. 29-33, 2004, doi: 10.1002/mop.11277.

[10] S. Chen and K. -M. Luk, "High performance dual-band dual-polarized magneto-electric dipole base station antenna," *Asia-Pacific Microwave Conference*, Sendai, Japan, 2014, pp. 321-323.

[11] A. Vedaee and H. R. Hassani, "A novel compact dual-band dual-polarized microstrip patch antenna for GSM/DCS applications," *Microwave and Optical Technology Letters*., vol. 58, no. 11, pp. 2557-2559, 2016, dio: 10.1002/mop.30099.

[12] Q. Hua, Y. Huang, A. Alieldin, C. Song, T. Jia, and X. Zhu, "A Dual-Band Dual-Polarized Base Station Antenna Using a Novel Feeding Structure for 5G Communications," in *IEEE Access*, vol. 8, pp. 63710-63717, 2020, doi: 10.1109/ACCESS.2020.2984199.

[13] H. Zhai, K. Zhang, S. Yang, and D. Feng, "A Low-Profile Dual-Band Dual-Polarized Antenna With an AMC Surface for WLAN Applications," in *IEEE Antennas and Wireless Propagation Letters*, vol. 16, pp. 2692-2695, 2017, doi: 10.1109/LAWP.2017.2741465.

[14] S. Khorasani, J. Nourinia, Ch. Ghobadi, A .Hatamian, and B. Virdee, "Dual- band magneto-electric dipole antenna with high-gain for base-station applications," *AEU - International Journal of Electronics and Communications*, vol. 134, p. 153696, May 2021, dio: 10.1016/j.aeue.2021.153696.

Declaration of Competing Interest: Authors do not have conflict of interest. The content of the paper is approved by the authors.

Author Contributions:
Maryam Mohammadifar: conducted the simulations and prepared the manuscript with input from all authors. **Pejman Mohammadi:** developed the theoretical framework and supervised the research. **Yashar Zehforoosh**: validated the analytical methods.

E-ISSN: 2980-9231
https://sanad.iau.ir/journal/jce

مهندسی مخابرات جنوب
دانشگاه آزاد اسلامی واحد بوشهر

https://doi.org/10.30495/jce.2025.1993480.1337

مقاله پژوهشی

یک آنتن BTS جدید برای کارکرد همزمان در باندهای فرکانسی GSM۹۰۰ و LTE

مریم محمدیفر[۱] | پژمان محمدی*[۲] | یاشار زهفروش[۳]

[۱] گروه برق، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه، ایران.
mohammadifar_eng@yahoo.com

[۲] مرکز تحقیقات مایکروویو و آنتن، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه، ایران.
P.mohammadi@iaurmia.ac.ir

[۳] مرکز تحقیقات مایکروویو و آنتن، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه ، ایران.
y.zehforoosh@srbiau.ac.ir

نویسنده مسئول:
* پژمان محمدی، دانشیار مرکز تحقیقات مایکروویو و آنتن، واحد ارومیه، دانشگاه آزاد اسلامی، ارومیه، ایران.
P.mohammadi@iaurmia.ac.ir

موضوع اصلی:
طراحی آنتن

تاریخچه مقاله:
تاریخ دریافت: ۴ مرداد ۱۴۰۳
تاریخ بازنگری: ۲۵ مرداد ۱۴۰۳
تاریخ پذیرش: ۱۳ شهریور ۱۴۰۳

چکیده:

در این مقاله، طراحی جدیدی از آنتن‌های دو بانده با پلاریزاسیون دوگانه برای ایستگاه‌های پایه موبایل جهت پشتیبانی از سیستم‌های ارتباطی سیار که در باندهای فرکانس (GSM/DCS/PCS/UMTS) سامانه جهانی ارتباطات سیار/ باند ذخیره جهت برقراری مکالمه/ سامانه ارتباط‌های شخصی/ سامانه جهانی مخابرات سیار و (LTE) نسل چهارم ارتباطات بیسیم کار می‌کنند، ارائه شده است. در این طراحی، به‌دلیل استفاده از تکنیک تغذیه سه‌گانه، یک پهنای باند گسترده با تطبیق امپدانسی عالی در ورودی آنتن به‌دست آمده است. دو دایپل چاپی که به‌صورت عمود بر یکدیگر قرار دارند و توسط خطوط میکرو استریپی پلکانی تغذیه می‌شوند، آنتن پیشنهادی را تشکیل می‌دهند. علاوه بر این، با قرار دادن یک ساختار جعبه‌ای ساده، به‌عنوان یک بازتابنده فلزی در زیر آنتن، تشعشع دو جهته دایپل‌ها با افزایش بهره آنتن به تشعشع یک‌طرفه تبدیل می‌شوند. نتایج طراحی این آنتن نشان می‌دهند که آنتن پیشنهادی مناسب برای کاربردهای ایستگاه پایه در فرکانس‌های عملیاتی ۹۰۰/۱۸۰۰/۱۹۰۰/۲۳۰۰ مگاهرتز می‌باشد. همچنین طبق نتایج بیشینه بهره ۱۱/۴۰ و ۱۰/۴۷ دسیبل به‌ترتیب در پورت‌های ۱ و ۲ آنتن به‌دست آمده است. ابعاد آنتن ۱۶۸×۱۶۸ میلی‌متر مربع بوده که بر روی یک بازتابنده جعبه‌ای شکل به ابعاد ۲۲۲ × ۲۲۲ میلی‌متر مربع با عمق ۴۲ میلی‌متر نصب شده است.

کلید واژه‌ها: آنتن‌های ایستگاه پایه موبایل، پلاریزاسیون دوگانه، آنتن‌های چاپی، تغذیه سه‌گانه

تازه‌های تحقیق:

- ساختار ساده آنتن نسبت به طراحی‌های گذشته
- طراحی جدیدی از آنتن‌های دو بانده با پلاریزاسیون دوگانه برای کارکرد همزمان در باندهای فرکانسی GSM۹۰۰ و LTE
- بهره خوب و بالا نسبت به طراحی‌های انجام شده گذشته

۱-مقدمه

طی سالهای گذشته رشد چشمگیری در زمینه ارتباطات بی‌سیم رخ داده است. حل مشکل محوشوندگی سیگنال در ارتباطات سیار یکی از مهم‌ترین چالش‌های طراحان آنتن بوده است. دلیل اصلی پدیده محوشوندگی سیگنال، پراکندگی و انعکاس امواج از اشیاء اطراف کاربران موبایل مانند ساختمان‌ها، اتومبیل‌ها و درختان است. بنابراین استفاده از آنتن‌های متعدد برای افزایش ظرفیت کانال و مهار محوشوندگی در این سیستم‌ها ضروری به نظر می‌رسد. از سوی دیگر، ایجاد کوپلینگ متقابل بین آنتن‌ها چالش دیگری برای سیستم‌های ارتباطی سیار است. یک راه حل معمول برای این مشکل استفاده از تنوع فضایی[1] (چندگانگی فضایی) است که به فضای بیشتری برای سیستم نیاز دارد [۱]. در نتیجه، در چند سال اخیر، انواع مختلفی از آنتن‌های ایستگاه پایه گزارش شده‌اند که ساختارهای مختلفی دارند و تنها از یک پلاریزاسیون پشتیبانی می‌کنند [۵]-[۲]. لازم به توضیح است که آنتن‌های مونوپل برای تکنیک‌های تنوع فضایی در سیستم‌های ایستگاه پایه مناسب هستند. در مرجع [۲]، یک آنتن ایستگاه پایه شامل یک تشعشع کننده دایپل پاپیونی که بر روی صفحه زمین قرار دارد گزارش شده است که در آن یک پل رسانا برای گسترش عرض بیم آنتن به‌کاررفته است. این آنتن فقط یک باند فرکانسی ۱۱/۵٪ (۶-۵/۳۵ گیگاهرتز) را پوشش می‌دهد. همچنین امروزه برای بهبود عملکرد ایستگاه‌های پایه، آنتن‌های ایستگاه پایه با عملکرد دو بانده پیشنهاد می‌شوند [۳]، [۵]. آنتن‌های دو بانده که از پچ‌های کوتاه و المان‌های مسطح با پهنای باند امپدانسی ۳۴٪ (۱/۱-۰/۷۸ گیگاهرتز) و ۴۹/۵٪ (۲/۶۲-۱/۵۸ گیگاهرتز) در [۳] و نیز ۱۰٪ (۲/۵-۲/۲۶ گیگاهرتز) و ۱۶ ٪ (۳/۸۸-۳/۲۸ گیگاهرتز) در [۴]، معرفی شده‌اند. در مرجع [۵] یک آنتن پچ میکرواستریپ دو بانده برای کاربردهای ایستگاه پایه فرستنده/گیرنده در باند فرکانس نسل چهارم ارتباطات بیسیم بررسی شده است. آنتن مذکور شامل دو دایپل یکسان و خطوط میکرواستریپ نامنظم است که از کابل کواکسیال برای تغذیه استفاده می‌کنند. اخیراً پیشرفت‌های صورت گرفته در زمینه سیستم‌های ارتباطی بی‌سیم و ارتقای دستگاه‌های ارتباطی سیار، نیازمند کاهش فضای نصب، هزینه تجهیزات و پشتیبانی از باندهای فرکانسی مختلف است. در نتیجه ارائه آنتن‌ها با عملکرد دو بانده و قابلیت پشتیبانی از دو پلاریزاسیون مختلف مورد نیاز است. در ابتدا، طراحی همزمان آنتن‌هایی که این ویژگی‌ها را دارند دشوار به نظر می‌رسید. چندین آنتن تک بانده با پلاریزاسیون دوگانه در مراجع [۸]-[۶] گزارش شده است. این آنتن‌ها عمدتاً از دو دایپل متعامد برای ایجاد پلاریزاسیون‌های دوگانه (خطی عمودی و خطی افقی) استفاده می‌کنند. با این وجود، آنتن‌های تک بانده نمی‌توانند تمام الزامات سیستم‌های ارتباطی مدرن را برآورده کنند. برای حل این مشکل، آنتن‌های دو بانده با پلاریزاسیون دوگانه برای کاربردهای ایستگاه پایه پیشنهاد می‌شوند [۱۳]-[۹]. یک آنتن ایستگاه پایه دو بانده با پلاریزاسیون دوگانه که باندهای سامانه جهانی ارتباطات سیار/ باند ذخیره جهت برقراری مکالمه/ سامانه ارتباط‌های شخصی/ سامانه جهانی مخابرات سیار را پوشش می‌دهد در [۸] ارائه شده است. این طراحی شامل چهار شکاف است که توسط خطوط میکرو استریپ در سمت بالا و چهار حفره در زیر شکاف‌ها تشکیل می‌شود. در این طرح دو شکاف برای ایجاد پلاریزاسیون خطی ۴۵+ درجه (اسلنت) و دو شکاف دیگر برای پلاریزاسیون خطی ۴۵- درجه (اسلنت) در نظر گرفته شده است. در این طراحی باندهای فرکانسی ۵/۶٪ (۰/۹۹۵-۰/۹۴۱ گیگاهرتز) و ۳۴/۹٪ (۲/۴۳۴-۱/۷۱ گیگاهرتز) برای پورت ۱ و همچنین باندهای فرکانسی ۶/۲٪ (۹۳۲/۹۹۲-۰/۰ گیگاهرتز) و ۱۹/۸٪ (۲/۰۷-۱/۷۰۲ گیگاهرتز) برای پورت ۲ به دست آمده‌اند.

یک آنتن دایپل دو بانده برای سیستم‌های بی‌سیم تلفن همراه مانند نسل دوم-نسل سوم و شبکه بیسیم محلی محبوب و پرکاربرد می‌باشد، زیرا پهنای باند امپدانسی آن می‌تواند دو باند فرکانسی مختلف را پوشش دهد و از دو پلاریزاسیون متعامد برای کاهش اثر محوشوندگی چند مسیری استفاده می‌نماید. ازاین‌رو در سال‌های اخیر طراحی‌های مختلفی از این آنتن‌ها پیشنهاد شده است. مفاهیم اساسی طراحی این آنتن‌ها را می‌توان به دو دسته کلی تقسیم کرد. دسته اول، طراحی‌هایی که بر اساس یک آنتن دو بانده با یک تغذیه است و عملیات پلاریزاسیون دوگانه با افزودن یک تغذیه متعامد به ساختار آنتن به دست می‌آید که نشان می‌دهد آنتن به‌صورت دو بانده با پلاریزاسیون دوگانه حاصل دو پورت ورودی می‌باشد. دسته دوم، طراحی‌های مبتنی بر یک ساختار انباشته متشکل از دو المان آنتن است که در فرکانس‌های مختلف کار می‌کنند و هر المان می‌تواند پلاریزاسیون‌های متعامد دوگانه را از طریق دو پورت ورودی ارائه دهد. در این طراحی، یک آنتن دو بانده با پلاریزاسیون دوگانه جدید معرفی شده

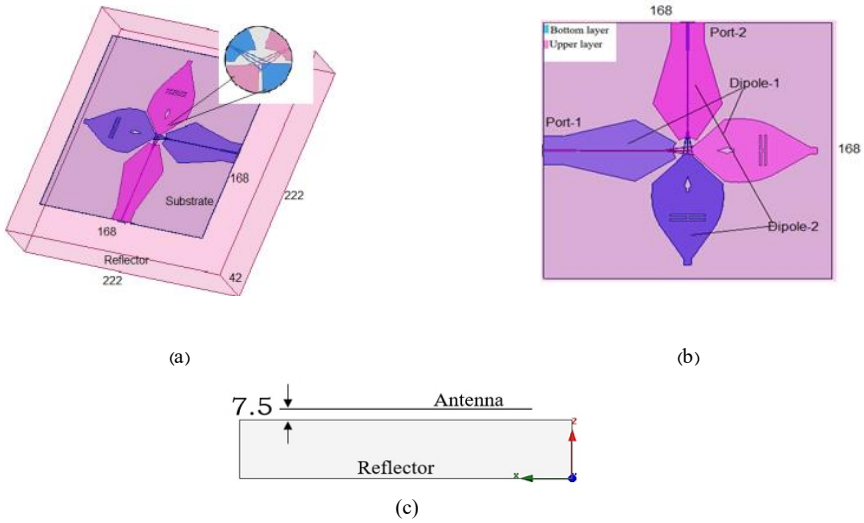
[1] Space Diversity

است که جهت افزایش پهنای باند فرکانسی آنتن و بهبود تطبیق امپدانسی آن از تکنیک تغذیه سه‌گانه در طراحی آن استفاده شده است. ساختار آنتن از دو دایپل چاپی که به‌صورت متعامد نسبت به هم قرار دارند تشکیل شده است بطوریکه دایپل‌ها توسط خطوط میکرو استریپی پلکانی تغذیه می‌شوند. علاوه بر این، با قرار دادن یک ساختار جعبه‌ای ساده، به‌عنوان یک بازتابنده فلزی در پشت دایپل‌ها، تشعشع دو جهته دایپل‌ها با افزایش بهره آنتن به تشعشع یک‌طرفه تبدیل می‌شوند. آنتن پیشنهادی با پوشش دو باند فرکانسی، برای ایستگاه‌های پایه موبایل جهت پشتیبانی از سیستم‌های ارتباطی سیار که در باندهای فرکانس سامانه جهانی ارتباطات سیار/ باند ذخیره جهت برقراری مکالمه/ سامانه ارتباط‌های شخصی/ سامانه جهانی مخابرات سیار و نسل چهارم ارتباطات بیسیم کار می‌کنند، مناسب می‌باشد. نتایج طراحی این آنتن نشان می‌دهند که آنتن دوبانده با پلاریزاسیون دوگانه پیشنهادی ، مناسب برای کاربردهای ایستگاه پایه در فرکانس‌های عملیاتی ۲۳۰۰/۱۹۰۰/۱۸۰۰/۹۰۰/۸۰۰ مگاهرتز می‌باشد. همچنین طبق نتایج بیشینه بهره ۱۰/۴۷ و ۱۱/۴۰ دسی بل در پورت‌های ۱ و ۲ آنتن به دست آمده است. ابعاد کلی آنتن ۱۶۸ × ۱۶۸ میلی‌متر مربع بوده که بر روی یک بازتابنده جعبه‌ای شکل به ابعاد ۲۲۲ × ۲۲۲ میلی‌متر مربع با عمق ۴۲ میلی‌متر نصب شده است.

۲- طراحی آنتن

آنتن‌های دایپل متعامد با پلاریزاسیون مایل دوگانه دسته‌ای از انواع مختلف آنتن‌ها هستند که معمولاً در کاربردهای ایستگاه‌های فرستنده پایه در حوزه مخابرات استفاده می‌شوند. این آنتن‌ها نقش مهمی در سیستم‌های ارتباطی بی‌سیم دارند و امکان انتقال و دریافت سیگنال‌ها را بین ایستگاه گیرنده و فرستنده و دستگاه‌های تلفن همراه فراهم می‌کنند. پیکربندی دایپل متعامد شامل دو المان دوقطبی است که عمود بر یکدیگر چیده شده‌اند و شکل متقاطع را تشکیل می‌دهند. این طراحی چند مزیت مهم از جمله پترن تشعشعی بهبود یافته و چندگانگی پلاریزاسیون را ارائه می‌دهد. پلاریزاسیون مایل دوگانه به استفاده از دو پلاریزاسیون مایل اشاره دارد که معمولاً در زاویه ۴۵ ± درجه از محورهای عمودی و افقی ایجاد می‌شوند. استفاده از پلاریزاسیون مایل دوگانه به کاهش اثرات محو شوندگی چند مسیری کمک می‌کند، پدیده‌ای که در آن سیگنال‌ها مسیرهای متعددی را برای رسیدن به گیرنده طی می‌کنند و منجر به اعوجاج و تخریب سیگنال می‌شود. با ارسال و دریافت سیگنال‌ها در هر دو جهت عمودی و افقی، آنتن دایپل متقاطع با پلاریزاسیون مایل دوگانه، پایداری و قابلیت اطمینان لینک ارتباطی را افزایش می‌دهد. در کاربردهای ایستگاه گیرنده و فرستنده، این آنتن‌ها به‌صورتی قرار می‌گیرند تا پوشش و ظرفیت بهینه را برای شبکه‌های ارتباطی سیار فراهم کنند. استفاده از پلاریزاسیون مایل دوگانه به کاهش محو شوندگی سیگنال کمک می‌کند و عملکرد آنتن را در محیط‌های مختلف بهبود می‌بخشد و آن را برای محیط‌های شهری، برون شهری و روستایی مناسب می‌کند. به‌طورکلی، آنتن‌های دایپل متعامد با پلاریزاسیون مایل دوگانه با توجه به چالش‌های مرتبط با انتشار و دریافت سیگنال در سناریوهای دنیای واقعی، به کارایی و کیفیت ارتباطات بی‌سیم در کاربردهای ایستگاه گیرنده و فرستنده کمک می‌کنند.

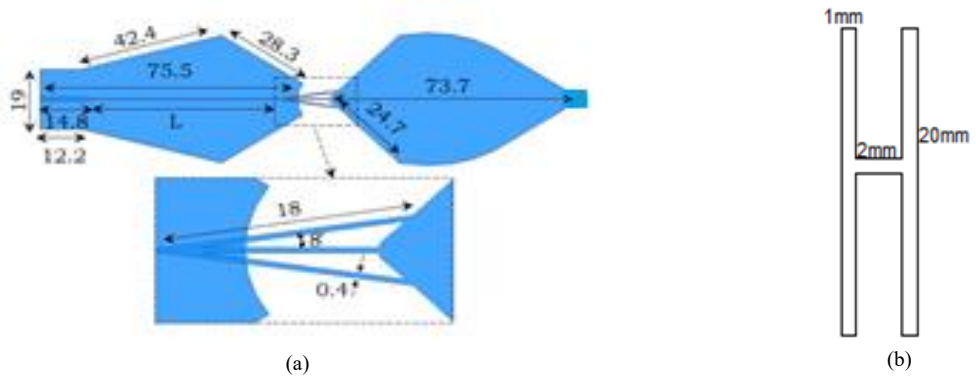
پیکربندی آنتن دایپل میکرواستریپ دو بانده پیشنهادی بدون شکاف در شکل ۱ نشان داده شده است. این آنتن از دو دایپل متعامد چاپی، دو ساختار تغذیه سه‌گانه برای هر تشعشع کننده و همچنین یک جعبه فلزی در زیر آنتن تشکیل شده است. در این طراحی از یک زیرلایه Rogers-RO4003 به ضخامت ۰/۵۰۸ میلی‌متر با گذردهی ۳/۵۵ و تانژانت تلفات ۰/۰۰۲۷ جهت چاپ دایپل ها در لایه‌های پایین و بالایی آن استفاده شده است. علاوه بر این، ساختار حفره‌ای منعکس‌کننده که از آلومینیوم ساخته شده است، در فاصله ۷/۵ میلی‌متری از زیرلایه قرار دارد. دایپل ها به‌طور مستقل با دو پورت تحریک می‌شوند. اساساً دایپل ها عمود بر یکدیگر هستند و در موقعیتی قرار دارند که پلاریزاسیون های مایل ۴۵+ و ۴۵- را ایجاد نمایند. نمای سه‌بعدی و نمای جانبی آنتن پیشنهادی در شکل ۱ نشان داده شده است.



شکل۱: پیکربندی آنتن پیشنهادی (a) نمای سه‌بعدی، (b) نمای بالا و (c) نمای جانبی آنتن پیشنهادی

Figure 1. Suggested antenna configuration: (a) 3D view, (b) Top view, (c) Side view

ابعاد آنتن و فاصله بازتابنده از آنتن در شکل نمایش داده شده است. برای جلوگیری از تماس دایپل ها با یکدیگر و بروز اتصال الکتریکی بین آن‌ها، دو بخش از هر دایپل در طرف‌های مختلف زیرلایه چاپ شده است. علاوه بر این، ابعاد دایپل پیشنهادی و ساختار تغذیه و همچنین اسلات H شکل ایجاد شده در روی دایپل ها در شکل‌های ۲ (a) و (b)، نمایش داده شده است. المان تشعشعی قطره‌ای شکل دارای طول ۷۳/۷ میلی‌متر و دو بازوی جانبی با طول‌های ۲۴/۷ میلی‌متر می‌باشد. با توجه به این شکل، دایپل با سه خط میکرو استریپ به نقطه تغذیه متصل می‌شود که به تغذیه سه‌گانه معروف است. یک صفحه زمین ذوزنقه‌ای شکل برای هر پورت از آنتن، ساختار دایپل ها را تکمیل می‌کند.

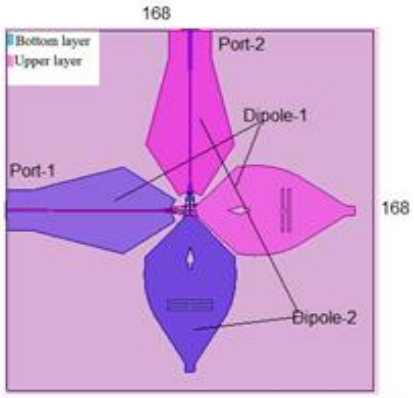


شکل۲: (a) ابعاد دایپل ها و ساختار تغذیه و (b) اسلات H شکل ایجاد شده در روی دایپل ها

Figure 2. (a) Dimensions of dipoles and feeding structure, (b) H-shaped slot created on the dipoles

هر کدام از دایپل های آنتن پیشنهادی توسط یک خط میکرو استریپ ۵۰ اهمی با عرض ۱/۱ میلی‌متر و طول ۱۴/۸ میلی‌متر تغذیه می‌شوند. خط میکرو استریپ دیگری با طول L (۵۴/۸ میلی‌متر) برای تطبیق امپدانسی بین ورودی و ساختار سه‌گانه استفاده شده است.

در نهایت برای افزایش بهره آنتن یک اسلات H شکل در طراحی ایجاد می‌کنیم و شکل نهایی با تغییرات به شکل زیر تغییر پیدا می‌کند.

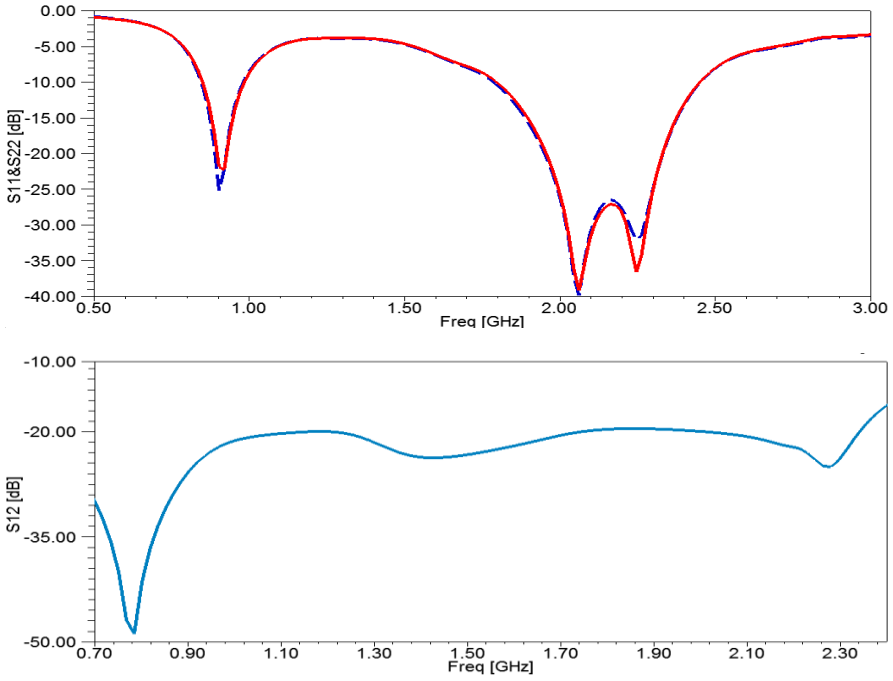


شکل ۳: شکل نهایی آنتن پیشنهادی

Figure 3. The final shape of the proposed antenna

۳- تحلیل و بررسی نتایج

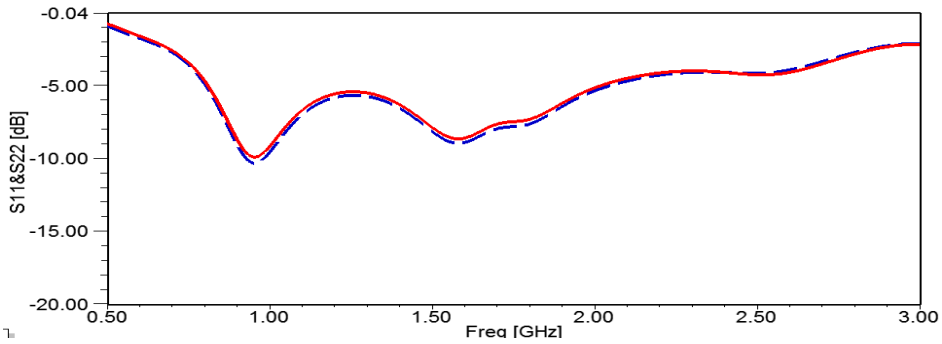
شکل ۴ نتایج شبیه‌سازی شده آنتن ارائه شده را نشان می‌دهد. تمامی شبیه‌سازی‌ها در این تحقیق با استفاده از نرم افزار ANSYS HFSS انجام شده است. با تحریک هر دو پورت آنتن پیشنهادی، پهنای باند امپدانس یکسانی برای پورت‌ها به‌دست می‌آید و باندهای فرکانسی ۸۵۲ تا ۹۸۶ مگاهرتز و ۱۷۶۰ تا ۲۴۶۳ مگاهرتز را به‌صورت مشابه پوشش می‌دهند. علاوه بر این، ایزولاسیون شبیه‌سازی شده بین پورت‌ها بالاتر از ۱۹ دسی بل است.



شکل۴: پارامترهای S شبیه‌سازی شده آنتن پیشنهادی

Figure 4. Simulated S parameters of the proposed antenna

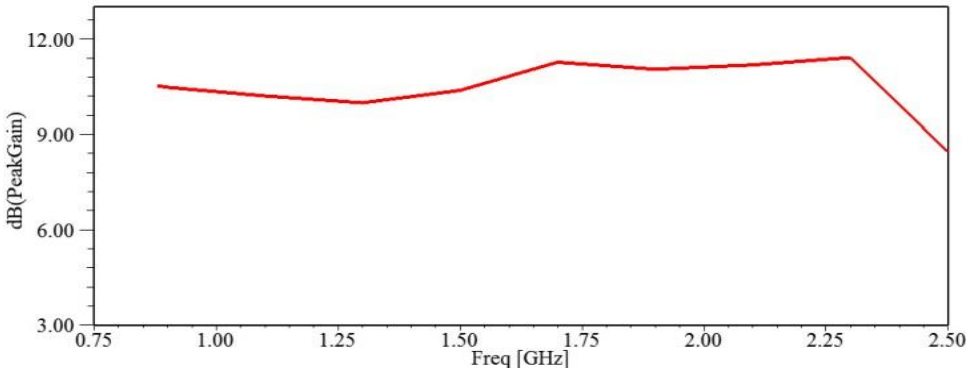
تغذیه میکرواستریپ سه‌گانه یکی از تکنیک‌هایی است که برای افزایش پهنای باند امپدانسی آنتن‌های چاپی استفاده می‌شود. پهنای باند امپدانسی یک پارامتر مهم است زیرا محدوده فرکانس‌هایی را که آنتن می‌تواند به‌طور مؤثر بر روی آن کار کند را تعیین می‌کند. تکنیک‌های مختلفی برای دستیابی به پهنای باند امپدانس وسیع‌تر استفاده می‌شود و تغذیه میکرو استریپی سه‌گانه یکی ازاین‌روش‌ها است. در این بخش این موضوع که چگونه تکنیک تغذیه سه‌گانه می‌تواند پهنای باند آنتن را افزایش دهد بررسی می‌شود. در واقع شکل سه‌گانه تغذیه میکرواستریپ به‌طور مؤثر طول الکتریکی ساختار تغذیه را افزایش می‌دهد. افزایش طول الکتریکی تغذیه، به آنتن اجازه می‌دهد تا در فرکانس‌های متعدد رزونانس کرده و در نتیجه پهنای باند را افزایش دهد. همچنین شکل سه‌گانه تغذیه، حالت‌های تشدید بیشتری را در مقایسه با یک آنتن میکرواستریپ ساده معرفی می‌کند. به‌این‌ترتیب، هر حالت رزونانسی به افزایش پهنای باند امپدانس کلی آنتن کمک می‌کند. از سوی دیگر، هندسه تغذیه میکرواستریپ سه‌گانه هر دو جزء القایی و خازنی را به ساختار تغذیه معرفی می‌کند. این مؤلفه‌ها می‌توانند به‌صورت استراتژیک برای دستیابی به تطبیق امپدانسی در فرکانس‌های متعدد طراحی شوند و پهنای باند را افزایش دهند. تغذیه میکرواستریپ سه‌گانه، توزیع جریان را در پچ تشعشعی اصلاح می‌کرده و این اصلاح می‌تواند منجر به توزیع یکنواخت‌تر جریان در محدوده فرکانسی وسیع‌تری شود که تطبیق امپدانسی و پهنای باند را بهبود می‌بخشد. خوراک سه‌تایی شکل می‌تواند طراحی یک شبکه تطبیق بهبود یافته را تسهیل کند. برای درک بهتر تأثیر تکنیک تغذیه سه‌گانه اعمال‌شده در طراحی پیشنهادی، افت بازگشتی آنتن پیشنهادی با یک تغذیه سه‌گانه و با یک تغذیه منفرد در شکل ۵ مقایسه شده است. نتایج نشان می‌دهند که ساختار تغذیه پیشنهادی به‌طور قابل توجهی پهنای باند آنتن را افزایش می‌دهد، به خصوص در باند فرکانس بالا. دلیل استفاده از ساختار تغذیه سه‌گانه در طرح پیشنهادی، پوشش باندهای فرکانسی عملیاتی BTS (۸۰۰/۹۰۰/۱۸۰۰/۱۹۰۰/۲۳۰۰ مگاهرتز) با تطبیق امپدانسی خوب می‌باشد.



شکل ۵: افت بازگشتی آنتن پیشنهادی با یک تغذیه منفرد

Figure 5. Return loss of the proposed antenna with a single feed

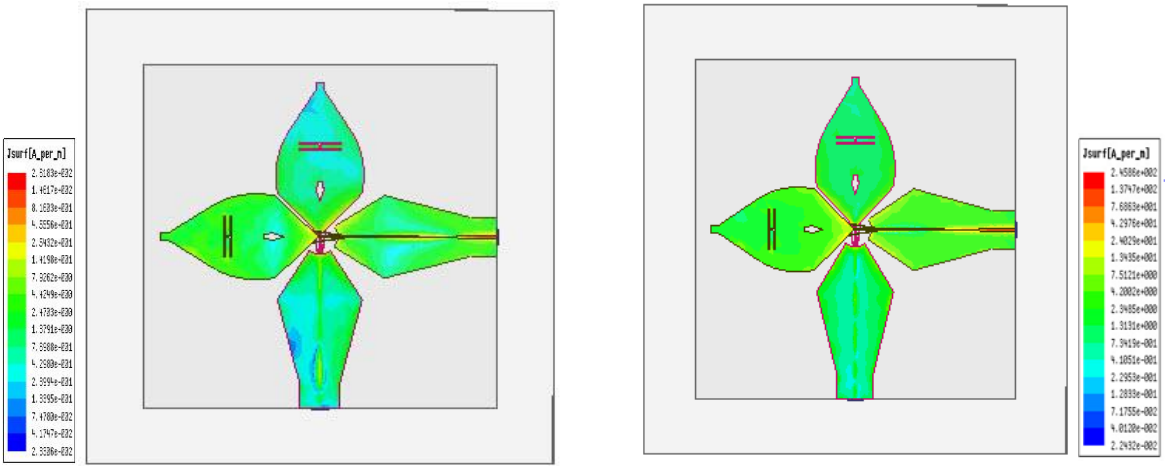
شایان ذکر است که تطبیق امپدانس و ایجاد رزونانس در آنتن‌های میکرواستریپ نه‌تنها به خط تغذیه بستگی دارد، بلکه تحت تأثیر شکل و اندازه صفحه زمین نیز قرار می‌گیرد. در این طراحی با همکاری خط میکرواستریپ باریک شونده، با تغذیه سه‌گانه و تنظیم طول و عرض شاخه‌های انتهایی تغذیه و سایر پارامترهای فیزیکی آنتن، تطبیق امپدانس عالی حاصل شده و در نتیجه پهنای باند امپدانسی وسیعی حاصل می‌شود. در واقع، با استفاده از ساختار تغذیه سه‌گانه برای تحریک المآن‌های تشعشعی آنتن، توزیع جریان عمودی روی صفحه تشعشع کننده‌ها به‌طور مؤثر افزایش می‌یابد، درحالی‌که توزیع جریان افقی کاهش می‌یابد و منجر به بهبود پهنای باند امپدانسی می‌شود. علاوه بر این، مقدار پیک بهره شبیه‌سازی شده آنتن پیشنهادی در شکل ۶ قابل مشاهده است.



شکل ۶: پیک بهره شبیه‌سازی شده آنتن

Figure 6. Simulated antenna gain peak

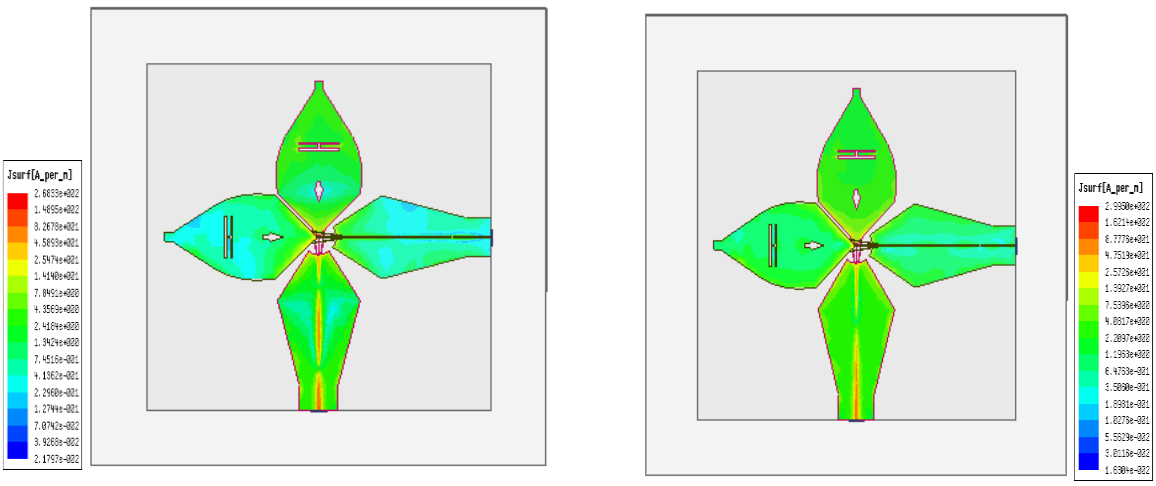
میدان‌های بازتاب شده در زیر زیرلایه در فرکانس‌های ۸۰۰ و ۲۳۰۰ مگاهرتز در شکل ۷ با تحریک پورت-۱ نشان داده شده است. با توجه به شکل، مشاهده می‌شود که تجمع میدان‌های الکتریکی در زیر خط تغذیه در فرکانس‌های پایین قوی‌تر از فرکانس‌های بالا است.



(a) (b)

شکل ۷: (a) توزیع میدان در فرکانس ۲۳۰۰ مگاهرتز برای پورت ۱، (b) توزیع میدان در فرکانس ۸۰۰ مگاهرتز برای پورت ۱

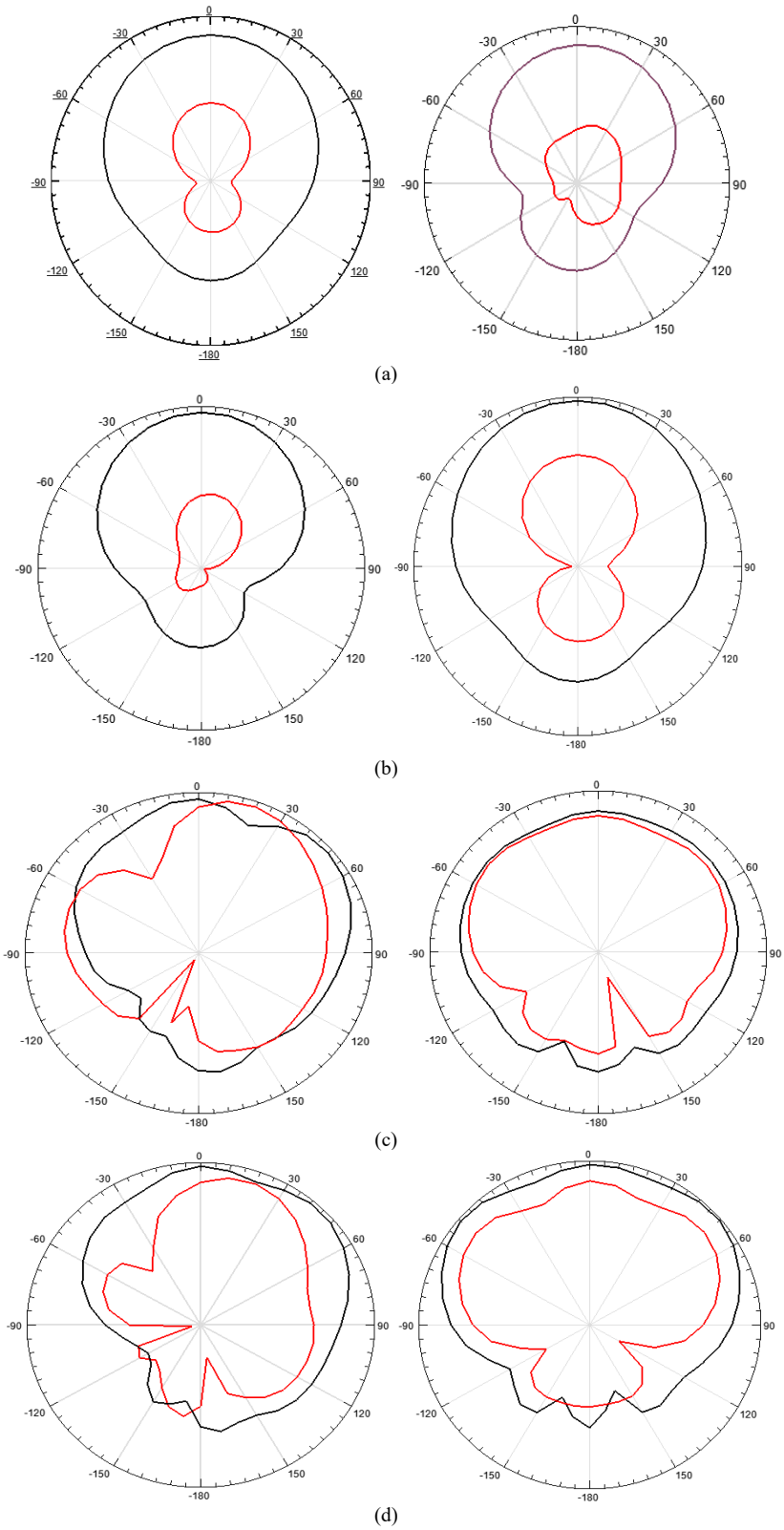
Figure7. (a) Field distribution at 2300MHz frequency for port1, (b) Field distribution at 800MHz frequency for port1



(a) (b)

شکل ۸: (a) توزیع میدان در فرکانس ۸۰۰ مگاهرتز برای پورت۲، (b) توزیع میدان در فرکانس ۲۳۰۰ مگاهرتز برای پورت ۲

Figure8. (a) Field distribution at 800MHz frequency for port2, (b) Field distribution at 2300MHz frequency for port2



شکل ۹: پترن‌های تشعشعی شبیه‌سازی شده برای پورت-۱ و پورت-۲. خطوط سیاه برای پلاریزاسیون Co و خطوط قرمز برای پلاریزاسیون Cross (a) پورت-۱ در ۹۰۰ مگاهرتز، (b) پورت-۲ در ۹۰۰ مگاهرتز، (c) پورت-۱ در ۲۳۰۰ مگاهرتز، (d) پورت-۲ در ۲۳۰۰ مگاهرتز

Figure 9. Simulated radiation patterns for port 1 and 2. Black lines for Co polarization and red lines for Cross polarization, (a) port 1 at 900MHz, (b) port 2 at 900MHz, (c) port 1 at 2300MHz, and (d) port 2 at 2300MHz

علاوه بر این، پترن های تشعشعی E-plane و H-plane آنتن در هر دو پورت در فرکانس‌های مختلف (۸۰۰ و ۲۳۰۰ مگاهرتز) در شکل ۹ نمایش داده شده است. شکل‌های ارائه شده نشان می‌دهند که آنتن دوبانده پیشنهادی دارای پترن تشعشعی بسیار پایدار در فرکانس‌های ذکر شده است. با توجه به پترن های تشعشعی آنتن پیشنهادی، مشاهده می‌شود که در فرکانس‌های بالاتر، مقداری اعوجاج در پترن‌های تشعشعی به وجود می‌آید که این موضوع می‌تواند به تحریک مودهای بالاتر در فرکانس‌های بالا مربوط باشد.

جدول ۱: مقایسه آنتن پیشنهادی با آنتن‌های مشابه

Table 1. Comparison of proposed antenna with similar antennas

Ref.	Feed	Gain Peak [dB]	Bandwidth [GHz]	Size [mm^3]	Structure
[12]	Double oval feeding shape	7/56 & 7/42	3/3 – 3/8 & 4/8 – 5/0	60 × 60 × 18	Complex
[14]	Γ-shape	9/4 7/9	13/31% (0/8- 0/91) 19/76% (1/7-2/08)	1/03 λ ×1/03 λ ×0/21 λ	Complex
[15]	Strip	6/7	40% (0/7–1/05) 60.87% (1/6–3)	220×220×100	Complex
[16]	Γ-shape	10/2	32.73% (0/69– 0/96) 45.45% (1/7–2/7)	360×360×31.5	Complex
This work	Microstrip	10/47 11/40	$\begin{cases}(0/852-0/986)\\(1/76-2/46)\end{cases}$	222×222×42	Simple

۳-نتیجه‌گیری

یک آنتن ایستگاه پایه جدید که از دایپل‌های متعامد، دو سری تغذیه سه‌گانه و یک ساختار منعکس‌کننده حفره‌ای تشکیل شده است، معرفی شده است. بررسی عملکرد آنتن پیشنهادی نشان می‌دهد که ویژگی دوگانگی پلاریزاسیون در دو باند فرکانسی به دست می‌آید و چندین باند فرکانسی مرتبط با ارتباطات تلفن همراه پوشش داده می‌شود. در طراحی ارائه شده، تکنیک تغذیه سه‌گانه اعمال‌شده، پهنای باند امپدانس آنتن را گسترش می‌دهد. همچنین با استفاده از ساختار منعکس‌کننده حفره‌ای شکل، یک پترن تشعشعی یک‌طرفه برای آنتن ایجاد می‌شود. پترن تشعشعی آنتن پایدار است و تغییر در مقدار پیک بهره در هر دو پورت کمتر از ۲ دسی بل است. در ساخت این آنتن از یک زیرلایه Rogers RO4003 با ابعاد ۱۶۸×۱۶۸ میلی‌متر مربع استفاده شده است. ساختار منعکس‌کننده حفره‌ای به ابعاد ۲۲۲×۲۲۲ میلی‌متر مربع و عمق ۴۲ میلی‌متر در زیر دایپل‌ها نصب شده است. طبق نتایج شبیه‌سازی ایزولاسیون بالای ۱۹ دسی بل و بیشینه بهره ۱۱/۴۰ و ۱۰/۴۷ دسی بل به ترتیب در پورت‌های ۱ و ۲ آنتن به دست آمده است. آنتن پیشنهادی با پوشش دو باند فرکانسی، برای ایستگاه‌های پایه موبایل جهت پشتیبانی از سیستم‌های ارتباطی سیار که در باندهای فرکانس سامانه جهانی ارتباطات سیار/ باند ذخیره جهت برقراری مکالمه/ سامانه ارتباط‌های شخصی/ سامانه جهانی مخابرات سیار و نسل چهارم ارتباطات بیسیم کار می‌کنند، مناسب می‌باشد.

مراجع

[1] B. Lindmark and M. Nilsson, “On the available diversity gain from different dual-polarized antennas,” *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 2, pp. 287-294, 2001, doi: 10.1109/49.914506.

[2] S. Qu, J. Li, and Q. Xue, “Bowtie Dipole Antenna with Wide Beamwidth for Base Station Application,” *IEEE Antennas and Wireless Propagation Letters*, vol. 6, pp. 293-295, 2007, doi: 10.1109/LAWP.2007.898543.

[3] W. X. An, H. Wong, K. L. Lau, S. F. Li, and Q. Xue, “Design of Broadband Dual-Band Dipole for Base Station Antenna,” *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 3, pp. 1592-1595, 2012, doi: 10.1109/TAP.2011.2180336.

[4] H. Zhai, Q. Gao, C. Liang, R. Yu, and S. Liu, “A Dual-Band High-Gain Base-Station Antenna for WLAN and WiMAX Applications,” *IEEE Antennas and Wireless Propagation Letters*, vol. 13, pp. 876-879, 2014, doi: 10.1109/LAWP.2014.2321503.

[5] M. Rezvani and P. Mohammadi, “Microstrip antenna with aperture reflector and C-shaped dipoles for LTE and wireless communications,” *Int J Electron Commun (AEÜ),* vol. 94, pp.12-18, 2018, doi: 10.1016/j.aeue.2018.06.041.

[6] Q. X. Chu, D. L. Wen, and Y. Luo, “A Broadband ± 45º Dual-Polarized Antenna With Y-Shaped Feeding Lines,” *IEEE Trans. Antennas Propag.*, vol. 63, no. 2, pp. 483-490, 2015, doi: 10.1109/TAP.2014.2381238.

[7] Y. Gou, S. Yang, J. Li, and Z. Nie, “A compact dual-polarized printed dipole antenna with high isolation for wideband base station applications,” *IEEE Trans. Antennas Propag*., vol. 62, no. 8, pp. 4392-4395, 2014, doi: 10.1109/TAP.2014.2327653.

[8] A. Elsherbini, J. Wu, and K. Sarabandi, “Dual polarized wideband directional coupled sectorial loop antennas for radar and mobile base-station applications,” *IEEE Trans. Antennas Propag*., vol. 63, no. 4, pp. 1505-1513, 2015, doi: 10.1109/TAP.2015.2392773.

[9] V. Deillon and J-F. Zürcher, and A. K. Skrivervik, “A compact dual-band dual-polarized antenna element for GSM/DCS/UMTS base stations,” *Microwave and Optical Technology Letters*., vol. 40, no. 1, pp. 29-33, 2004, doi: 10.1002/mop.11277.

[10] S. Chen and K. -M. Luk, "High performance dual-band dual-polarized magneto-electric dipole base station antenna," *Asia-Pacific Microwave Conference*, Sendai, Japan, 2014, pp. 321-323.

[11] A. Vedaee and H. R. Hassani, “A novel compact dual-band dual-polarized microstrip patch antenna for GSM/DCS applications,” *Microwave and Optical Technology Letters*., vol. 58, no. 11, pp. 2557-2559, 2016, dio: 10.1002/mop.30099.

[12] Q. Hua, Y. Huang, A. Alieldin, C. Song, T. Jia, and X. Zhu, "A Dual-Band Dual-Polarized Base Station Antenna Using a Novel Feeding Structure for 5G Communications," in *IEEE Access*, vol. 8, pp. 63710-63717, 2020, doi: 10.1109/ACCESS.2020.2984199.

[13] H. Zhai, K. Zhang, S. Yang, and D. Feng, "A Low-Profile Dual-Band Dual-Polarized Antenna With an AMC Surface for WLAN Applications," in *IEEE Antennas and Wireless Propagation Letters*, vol. 16, pp. 2692-2695, 2017, doi: 10.1109/LAWP.2017.2741465.

[14] S. Khorasani, J. Nourinia, Ch. Ghobadi, A .Hatamian, and B. Virdee, “Dual- band magneto-electric dipole antenna with high-gain for base-station applications,” *AEU - International Journal of Electronics and Communications*, vol. 134, p. 153696, May 2021, dio: 10.1016/j.aeue.2021.153696.

[15] A. Alieldin, Y. Huang, S. J. Boyes, M. Stanley, S. D. Joseph, and B. Al-Juboori, "A Dual-Broadband Dual-Polarized Fylfot-Shaped Antenna for Mobile Base Stations Using MIMO Over-Lapped Antenna Subarrays," in *IEEE Access*, vol. 6, pp. 50260-50271, 2018, doi: 10.1109/ACCESS.2018.2868817.

[16] H. Huang, X. Li, and Y. Liu, “A dual-broadband base station antenna with ikebana-like arrangement scheme,” *Microw Opt Technol Let.*, vol. 62, no. 2, pp. 708-713, 2019, doi: 10.1002/mop.32050.

JOURNAL OF SOUTHERN COMMUNICATION ENGINEERING
ISLAMIC AZAD UNIVERSITY BUSHEHR BRANCH

E-ISSN: 2980-9231
https://sanad.iau.ir/journal/jce

https://doi.org/10.30495/jce.2025.1993480.1334

Vol. 14/ No. 56/2025

Review Article

Software Project Scheduling Problem: A Review

Javad Pashaei Barbin [1*] | Mahdi Jalali[2]

[1]Department of Computer Engineering, Nag.C., Islamic Azad University, Naghadeh, Iran
Javad.pashaei@iau.ac.ir

[2]Department of Electrical Engineering, Nag.C., Islamic Azad University, Naghadeh, Iran
Mahdi.jalali@iau.ac.ir

Correspondence
Javad Pashaei Barbin, Assistant Professor of Computer Engineering, Nag.C., Islamic Azad University, Naghadeh, Iran.
Javad.pashaei@iau.ac.ir

Main Subjects:
Software Projects Scheduling

Paper History:
Received: 23 April 2024
Revised: 24 June 2024
Accepted: 26 June 2024

Abstract

The software project scheduling problem (SPSP) is one of the most important activities in software project development. One of the main reasons for unsuccessful software project completion is non-compliance with cost and schedule plans, often due to inefficient scheduling methods. The key factor for delivering software projects within planned cost and schedule is employing accurate and correct scheduling. SPSP is the most critical issue in project development and management, requiring more attention than any other aspect. Software project development should be fundamentally based on it. SPSP encompasses resource planning, cost estimation, manpower allocation, and cost control. Therefore, adopting an algorithm for software project scheduling that optimizes project completion time while considering cost and resource constraints is essential. Simultaneously reducing both cost and time in software project development is crucial for software production companies. Achieving a balance between project time and cost is necessary to minimize the asymmetry between these two factors. In SPSP, the most important element is the Resource Constrained Project Scheduling Problem (RCPSP). RCPSP involves assigning multiple tasks to limited-capacity resources under time constraints to optimize task scheduling with minimal time while satisfying and optimizing resource limitations. This article reviews SPSP using classical models and artificial intelligence algorithms.

Keywords: Software project scheduling problem, Project scheduling problem with limited resources, Heuristic algorithms.

Highlights

- A comprehensive and structured classification of software project scheduling methods.
- A detailed comparative analysis of algorithm performance across various scenarios.
- Identification of research gaps and future directions.

Citation: J. Pashaei Barbin, and M.Jalali, "Software Project Scheduling Problem: A Review," *Journal of Southern Communication Engineering*, vol. 14, no. 56, pp. 89-117, 2025, doi:10.30495/jce.2025.1993480.1334 [in Persian].

1. Introduction

The Software Project Scheduling Problem (SPSP) represents a fundamental challenge in software project management. Research indicates that inefficient scheduling constitutes one of the primary causes of software project failure [1]. Consequently, implementing effective models and algorithms for task and resource scheduling proves critical for project success.

Software project scheduling encompasses planning, organization, team supervision, and executive leadership activities, aiming to deliver specified outcomes within agreed costs and timelines through optimal resource utilization [2]. While initial estimates may appear optimistic for technically advanced projects with comprehensive planning, scheduling essentially decomposes the project into discrete activities with time estimates. Given that many activities execute concurrently, schedulers must coordinate these parallel tasks to optimize manpower and resource allocation [3].

Proper scheduling serves as a cornerstone for successful software project management. Current scheduling methodologies include:

- Critical Path Method (CPM)
- Program Evaluation and Review Technique (PERT)
- Graphical Evaluation and Review Technique (GERT) [3]

However, these traditional methods face limitations in activity duration estimation, leading to challenges in modeling real-world projects. Artificial intelligence approaches offer promising solutions to address these inherent limitations.

2. Objective

The main objective of this review paper is to present a comprehensive classification and comparative analysis of the methods and algorithms used to address SPSP. The focus is on classical models, metaheuristic algorithms, and machine learning techniques. We have organized the overall structure of the article as follows:

- In the second section, we will address the software project-scheduling problem
- In the third section, we will discuss different models of the software project-scheduling problem
- In the fourth section, we will discuss conclusions and future work.

3. Methodology

First, classical models such as CPM, PERT, GERT, and PDM are reviewed [4]. Then, a wide range of metaheuristic algorithms, including Genetic Algorithm [5], Ant Colony Optimization [6], Particle Swarm Optimization [7], Firefly Algorithm [8], and hybrid approaches [9] are analyzed. Additionally, the role of machine learning, particularly Artificial Neural Networks (ANNs), in time and cost estimation is discussed [10].

4. Key Findings

The results demonstrate that metaheuristic algorithms, especially hybrid approaches, outperform classical models in managing the uncertainty and complexity of real-world software projects [11],[12]. Machine learning methods further enhance scheduling accuracy by leveraging historical data patterns [10]. Our analysis confirms that artificial intelligence techniques exhibit superior performance compared to classical scheduling approaches. The complete classification of algorithms addressing software project scheduling is presented in Figure I.

5. Conclusion

This study demonstrates that intelligent algorithms, particularly hybrid metaheuristics, offer superior effectiveness for solving SPSP in practical implementations. Furthermore, incorporating machine learning techniques enhances decision-making capabilities for resource allocation and timeline estimation. We propose a hybrid algorithm for software project scheduling optimization that integrates genetic algorithms with neighborhood search. This combined approach achieves faster convergence while minimizing the risk of local optima. Comparative analysis reveals the algorithm's competitive efficiency against standard genetic and tabu search methods. Table I presents an evaluation of various metaheuristic and hybrid algorithms against software project scheduling criteria, using benchmark data from the PSPLIB database and j30 sample set.

6. Acknowledgement

The author would like to express sincere appreciation to the Islamic Azad University, Naqhadeh Branch, for providing the academic environment and support necessary for the completion of this research. Special thanks are also extended to the faculty members and students of the Computer Engineering Department for their constructive feedback and encouragement throughout the study.

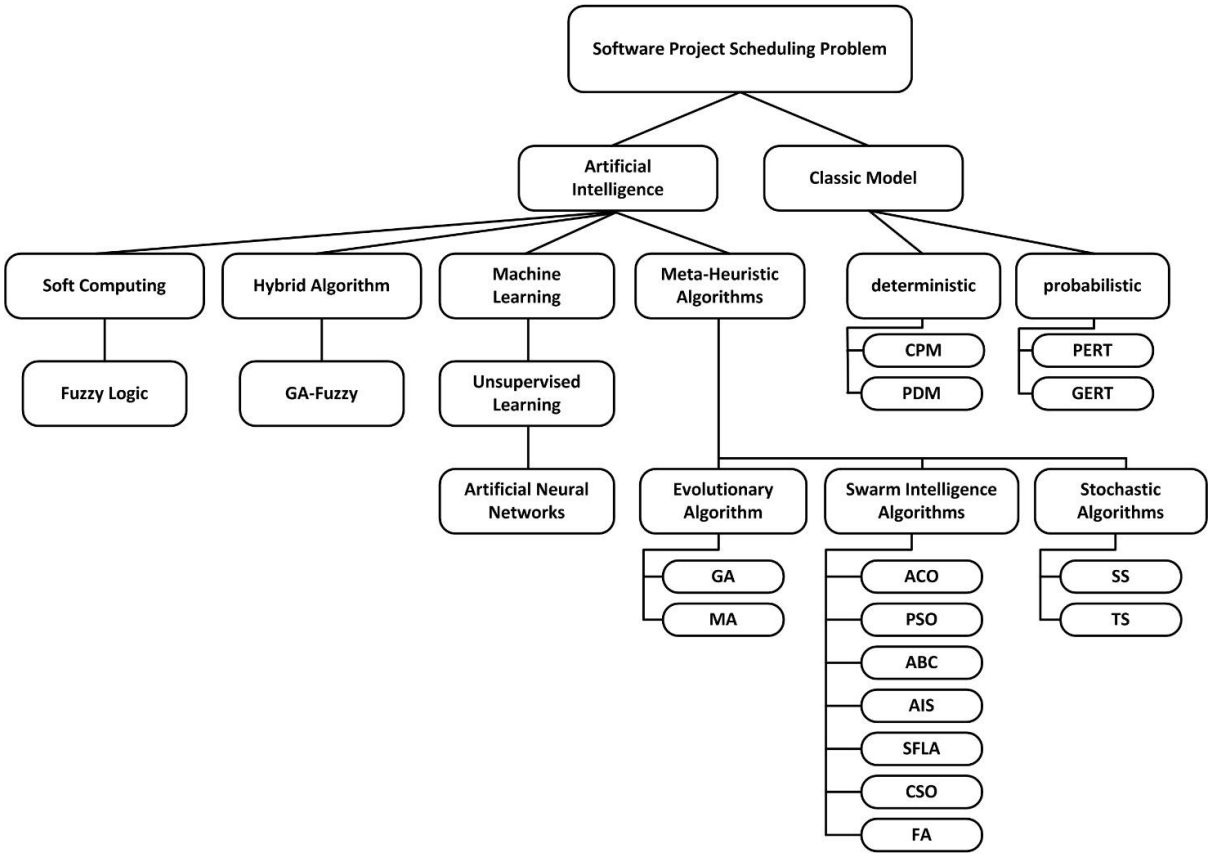


Figure I. Classification of different models in SPSP

Table I. Comparison of metaheuristic and hybrid algorithms in terms of project scheduling criteria

Maximum computation time (seconds)	Average computation time (seconds)	Mean of the standard deviation from the critical path	Mean percentage deviations from the desired limit	Algorithm
5.5	1.17	11.45	0.15	Genetic Algorithm
4.17	1.15	12.75	0.35	Bee colony
NR	NR	11.73	0.27	Scatter search
3.2	NR	11.71	0.46	Tabu search
NR	0.76	11.94	0.22	Hybrid

References

[1] W. Herroelen, E. Demeulemeester, and B.D. Reyck, "Resource Constrained Project Scheduling: A Survey of Recent Developments," *Computers & Operations Research*, vol. 25, no. 4, pp. 279-302, 1998, doi: 10.1016/S0305-0548(97)00055-5.

[2] E. Demeulemeester, "Minimizing Resource Availability Costs in Time-limited Project Networks," *Management Science*, vol. 41, pp. 1590-1598, 1995, doi: 10.1287/mnsc.41.10.1590.

[3] J. Zhang, X. Shen, and C. Yao, "Evolutionary Algorithm for Software Project Scheduling Considering Team Relationships," in *IEEE Access*, vol. 11, pp. 43690-43706, 2023, doi: 10.1109/ACCESS.2023.3270163.

[4] L.D. Long and A. Ohsato, "Fuzzy critical chain method for project scheduling under resource constraints and uncertainty," *International Journal of Project Management*, vol. 26, no. 6, pp. 688-698, 2008, doi: 10.1016/j.ijproman.2007.09.012.

[5] J. Holland, "Adaptation in Natural and Artificial Systems," University of Michigan, Michigan, USA, April 29, 1992.

[6] M. Dorigo and L. M.Gambardella, "The Colony System: A Cooperative Learning Approach to the Traveling Salesman Problem," *IEEE Transactions on Evolutionary Computation*, vol. 1, no.1, April 1997, doi: 10.1109/4235.585892.

[7] J. Kennedy and R. C. Eberhart, "Particle Swarm Optimization," in *Proceedings of the IEEE International Conference on Neural Networks*, 1995, pp. 1942-1948, doi: 10.1109/ICNN.1995.488968.

[8] X.S. Yang, Nature-Inspired Meta-heuristic Algorithms. Luniver Press, 2008.

[9] J. Lin, L. Zhu and K. Gao, "A genetic programming hyper-heuristic approach for the multi-skill resource constrained project scheduling problem," *Expert Systems with Applications*, vol. 140, p. 112915, 2020, doi: 10.1016/j.eswa.2019.112915.

[10] A. Agarwal, V.S. Jacob, and H. Pirkul, "Augmented Neural Networks for Task Scheduling," *European Journal of Operational Research*, vol. 151, no. 3, pp. 481-502, 2003, doi: 10.1016/S0377-2217(02)00605-7.

[11] Y. j. Shi, F.Z. Qu, W. Chen, and B. Li, "An Artificial Bee Colony with Random Key for Resource-Constrained Project Scheduling, *" Life System Modeling and Intelligent Computing. ICSEE LSMS 2010. Lecture Notes in Computer Science*, vol. 6329, 2010, pp. 148-157, doi: 10.1007/978-3-642-15597-0_17.

[12] M. Dam and M. Zachariasen, "Tabu Search on the Geometric Traveling Salesman Problem," *Meta-Heuristics: Theory and Applications*, pp. 571-587, 1995, doi: 10.1007/978-1-4613-1361-8_34.

Declaration of Competing Interest: Authors do not have conflict of interest. The content of the paper is approved by the authors.

Author Contributions:
Javad Pashaei Barbin: literature review, data analysis, and manuscript writing; **Mahdi Jalali**: methodology, manuscript editing.

E-ISSN: 2980-9231
https://sanad.iau.ir/journal/jce

مهندسی مخابرات جنوب
دانشگاه آزاد اسلامی واحد بوشهر

https://doi.org/10.30495/jce.2025.1993480.1334

مقاله مروری

مروری بر مساله زمانبندی پروژه‌های نرم‌افزاری

جواد پاشائی باربین* [1] | مهدی جلالی [2]

[1] گروه کامپیوتر، واحد نقده، دانشگاه آزاد اسلامی، نقده، ایران
javad.pashaei@iau.ac.ir

[2] گروه برق، واحد نقده، دانشگاه آزاد اسلامی، نقده، ایران
Mahdi.jalali@iau.ac.ir

نویسنده مسئول
* جواد پاشائی باربین ، استادیار، گروه کامپیوتر ، واحد نقده، دانشگاه آزاد اسلامی، نقده، ایران
javad.pashaei@iau.ac.ir

موضوع اصلی:
زمانبندی پروژه‌های نرم‌افزاری

تاریخچه مقاله:
تاریخ دریافت: ۴ اردیبهشت ۱۴۰۳
تاریخ بازنگری: ۴ تیر ۱۴۰۳
تاریخ پذیرش: ۶ تیر ۱۴۰۳

چکیده:

مساله زمانبندی پروژه‌های نرم‌افزاری یکی از مهمترین فعالیت‌ها در توسعه پروژه‌های نرم‌افزاری به‌شمار می‌آید. یکی از اصلی‌ترین دلایل عدم اتمام موفقیت‌آمیز پروژه‌های نرم‌افزاری، مطابقت نداشتن هزینه و زمانبندی برنامه‌ریزی شده و استفاده نکردن از روش‌های زمانبندی کارا است. عامل اصلی برای به‌اتمام رساندن پروژه‌های نرم‌افزاری مطابق با هزینه و زمانبندی برنامه‌ریزی شده، بکارگیری زمانبندی دقیق و درست است. موضوع زمانبندی پروژه‌های نرم‌افزاری، مهمترین مبحثی است که در توسعه و مدیریت پروژه می‌بایست بیش از هر موضوع دیگری به آن توجه کرد. این مساله شامل برنامه‌ریزی منابع، برآورد هزینه‌ها، نیروی انسانی و کنترل هزینه است. بنابراین، لازم است برای زمانبندی پروژه‌های نرم‌افزاری الگوریتمی اتخاذ شود که با در نظر گرفتن هزینه و محدودیت‌های منابع، بهینه‌ترین زمان برای انجام پروژه‌ها پیش‌بینی گردد. کاهش همزمان هزینه و زمان در توسعه پروژه‌های نرم‌افزاری برای شرکت‌های تولید نرم‌افزار بسیار لازم و ضروری است. بنابراین، به‌دلیل کاهش عدم تقارن دو عامل ذکر شده در پروژه‌ها لازم است که بین زمان پروژه و هزینه موازنه‌ای صورت پذیرد. در این مقاله به بررسی و مرور مساله زمانبندی پروژه‌های نرم‌افزاری با استفاده از مدل‌های کلاسیک و الگوریتم‌های هوش مصنوعی پرداخته شده است. بررسی‌های به دست آمده نشان می‌دهد که روش‌های هوش مصنوعی کارایی بهتری در مقایسه با مدل‌های کلاسیک دارند.

کلید واژه‌ها: مساله زمانبندی پروژه‌های نرم‌افزاری، مساله زمانبندی پروژه با منابع محدود، الگوریتم‌های فرااکتشافی

تازه های تحقیق:

- طبقه‌بندی جامع و ساخت‌یافته از روش‌های زمان‌بندی پروژه‌های نرم‌افزاری.
- تحلیل تطبیقی دقیق عملکرد الگوریتم‌ها در سناریوهای مختلف.
- شناسایی خلأهای تحقیقاتی و جهت‌گیری‌های آینده.

۱-مقدمه

مسئله‌ی زمانبندی پروژه‌های نرم‌افزاری[1] از مسائل غیر چند جمله‌ای سخت[2] است که به تعیین زمانبندی مناسب برای تخصیص منابع به وظایف می‌پردازد [۱]. در SPSP هدف اصلی یافتن راه حلی برای تخصیص منابع به وظایف است به گونه‌ای که هزینه و زمان پروژه‌های نرم‌افزاری حداقل گردند. بنابراین، هدف SPSP، این است که پروژه‌ها در قالب هزینه و زمانی معین و از پیش تعیین شده به اتمام برسند [۲]. هدف SPSP مشخص نمودن محدوده پروژه، تخمین حجم کاری مورد نیاز و در نتیجه تهیه یک برنامه زمانبندی جهت اجرای پروژه است.

زمانبندی پروژه با شناسایی نیازمندیهای پروژه آغاز می‌شود. نظارت بر پروژه و کنترل آن بر اساس معیارهای مهندسی نرم افزار تضمین می‌کند که روند اجرای وظایف در تیم‌ها و تخصیص منابع به آن‌ها بر اساس طرح اجرایی پروژه انجام پذیرد و مدیر پروژه در صورت مشاهده عدم تطابق در روند اجرا، اقدامات اصلاحی را به عمل آورد. با مساله زمانبندی پروژه‌های نرم‌افزاری، مدیر پروژه در طول چرخه توسعه نرم‌افزار، از پشتوانه قدرتمندی جهت اتخاذ تصمیمات مختلف برخوردار است. همچنین مدیر پروژه، تحلیل‌گر، طراح، برنامه‌نویس و سایر افراد توسعه نرم‌افزار باید بدانند که برای تولید یک پروژه نرم‌افزاری به چه میزان هزینه و زمان نیاز دارند. بدون داشتن یک برنامه زمانبندی مناسب از هزینه و زمان مورد نیاز برای توسعه پروژه‌های نرم‌افزاری، مدیر پروژه نمی‌تواند تشخیص دهد که به چه میزان زمان و چه مقدار هزینه جهت انجام پروژه نیاز دارد و در صورت اشتباه، پروژه در مسیر شکست حتمی یا با ریسک شکست مواجه می‌شود [۳].

مساله زمانبندی پروژه‌های نرم‌افزاری شامل فعالیت‌های برنامه‌ریزی، سازماندهی، نظارت بر تیم و هدایت اجرایی تیم را در بر می‌گیرد و سعی دارد تا با استفاده درست از منابع، نتایج مشخص و مورد انتظار را با هزینه توافق شده در موعد دقیق تحویل دهد [۴].مدیر پروژه‌های نرم‌افزاری باید زمان و منابع مورد نیاز برای انجام فعالیت‌ها را برآورد نماید و آنها را به ترتیب منسجمی سازمان‌دهی نماید. در زمانبندی پروژه‌های نرم‌افزاری حتی اگر پروژه‌های جدید شبیه به پروژه‌های قبلی باشند، زمانبندی پروژه‌های قبلی نمی‌تواند مبنای درستی برای پروژه‌های جدید باشند. با توجه به اینکه پروژه‌های نرم‌افزاری مختلف ممکن است از روش‌های مختلفی برای طراحی و پیاده‌سازی استفاده کنند لذا زمانبندی پیچیده‌تر خواهد شد.

اگر پروژه‌ها از نظر تکنیکی پیشرفته باشد و مدیران تمام جنبه‌ها را در نظر گرفته باشند، برآوردهای اولیه می‌تواند خوشبینانه باشد. زمانبندی، کل پروژه را به فعالیتهای جداگانه‌ای تقسیم می‌کند و زمان مورد نیاز برای کامل کردن این فعالیتها را برآورد می‌کند. معمولاً بعضی از فعالیتها به صورت موازی انجام می‌شوند. لذا زمانبندی باید این فعالیت‌های موازی را هماهنگ کنند و کار را طوری سازماندهی کند که نیروی انسانی و منابع به طور بهینه مورد استفاده قرار گیرند [۵]. در پروژه‌های نرم‌افزاری نباید وضعیتی به وجود آید که کل پروژه به خاطر عدم اتمام منابع، به تاخیر بیافتد. در زمانبندی، مدیران پروژه‌های نرم‌افزاری نباید فرض کنند که هر مرحله از پروژه فاقد هر گونه مشکلی است. اگر پروژه‌ها جدید باشند و از نظر تکنیکی پیشرفته باشند، بعضی از بخش‌های آن ممکن است دشوار باشند و ممکن است زمان بیشتری از زمان تخمین زده را به خود اختصاص دهند.

زمانبندی صحیح یکی از ارکان اصلی و لازمه موفقیت در مدیریت پروژه‌های نرم‌افزاری است. روش‌های متعددی برای زمانبندی پروژه‌ها وجود دارند که ازجمله آنها می‌توان به روش مسیر بحرانی[3] ، روش ارزیابی و بررسی برنامه[4] و روش ارزیابی و بررسی گرافیکی[5]اشاره کرد [۵]. این روش‌ها، به علت ضعف‌های موجود در نحوه تخمین مدت زمان فعالیت‌ها، برای مدل کردن پروژه‌های واقعی با مشکلات متعددی مواجه می‌شوند. یکی از راهکارهای اساسی در برخورد با چنین مشکلاتی استفاده از روش‌های هوش مصنوعی است.

ساختار کلی مقاله را به شرح زیر سازماندهی کرده‌ایم: در بخش دوم، به مساله زمانبندی پروژه‌های نرم‌افزاری خواهیم پرداخت؛ در بخش سوم، مدلهای مختلف مساله زمانبندی پروژه‌های نرم‌افزاری را بحث خواهیم کرد و در نهایت در بخش چهارم به نتیجه‌گیری و کارهای آینده می‌پردازیم.

---

[1] Software Project Scheduling Problem (SPSP)
[2] Non-Deterministic Polynomial (NP-Hard)
[3] Critical Path Method (CPM)
[4] Program Evaluation and Review Technique (PERT)
[5] Graphical Evaluation & Review Technique (GERT)

۲- مساله زمانبندی پروژه‌های نرم‌افزاری

زمانبندی عبارت است از تعیین یک توالی زمانی جهت انجام یک سری فعالیت‌های وابسته که تشکیل دهنده پروژه هستند [۶]. وابستگی فعالیت‌ها از نظر تقدم و تاخر بسیار مهم است. یعنی ممکن است انجام یک فعالیت به انجام چند فعالیت دیگر وابسته باشد که در این صورت گفته می‌شود که پروژه دارای محدودیت‌های تقدمی است. تعیین این برنامه زمانبندی با در نظر گرفتن هدف یا اهداف خاصی صورت می‌گیرد. تقریبا در تمامی پروژه ها محدودیتهای تقدمی بین فعالیت‌ها وجود دارند اما علاوه بر این محدودیت‌ها، ممکن است نوع دیگری از محدودیت‌ها تحت عنوان محدودیت‌های منابع نیز در پروژه وجود داشته باشد. بنابراین در زمانبندی پروژه علاوه بر اینکه باید به محدودیت‌های تقدمی توجه داشت، برنامه‌ریزی باید به گونه‌ای انجام شود که با محدودیت‌های منابع نیز سازگار باشد. در SPSP در نظر گرفتن مجموعه‌ای از اعمال برای هر فعالیت الزامی است. مهمترین فاکتورهای تاثیرگذار در SPSP در جدول ۱ نمایش داده شده است.

جدول ۱: پارامترهای تاثیرگذار در SPSP

Table 1. Effective parameters in SPSP

Effective parameters	Discriptions
$SK = \{s_1, s_2, ..., s_{\|SK\|}\}$	A set of skills required in project management
$TK = \{t_1, t_2, ..., t_{\|TK\|}\}$	Task set
$EM = \{e_1, e_2, ..., e_{\|EM\|}\}$	Employee set
V= TK	A set of vertices (tasks) in project scheduling
$A = \{(t_1, t_2), (t_3, t_4), ...(t_m, t_n)\}$	A set of arcs (priority relationships between tasks)
G(V,A)	Task precedence graph
t_j^{effort}	Amount of time for task tj (in people per month)
$t_j^{skills} \subseteq SK$	A set of skills required for task t_j
$e_i^{skills} \subseteq SK$	A set of skills for employee e_i
e_i^{salary}	A monthly salary for employee e_i
e_i^{Maxed}	Maximum time allocated by employee ei to task t_j
t_j^{start}	Task start time t_j
t_j^{end}	Task end time t_j
t_j^{cost}	Task cost amount t_j
$t_j^{duration}$	Task duration t_j
$p^{duration}$	Total duration of software project
p^{Cost}	Total cost of software project
p^{Over}	Total overtime of software project
e_i^{Over}	Employee overtime ei
$M = (m_{ij})_{E \times T}$	Solution matrix for software project scheduling problem

مساله زمانبندی پروژه‌های نرم‌افزاری یک گام مهم در فرایند توسعه نرم‌افزار است و از آن برای تحلیل در مورد منابع زمان و هدایت کلی پروژه استفاده می‌شود. هدف اصلی تخصیص منابع به فعالیت‌ها و زیر فعالیت‌های پروژه است که این منابع می‌توانند منابع نرم‌افزاری سخت‌افزاری و انسانی باشند. هدف تخصیص بهینه این منابع است رعایت به‌طوریکه با رعایت روابط

تقدمی در بین فعالیت‌ها زمان تکمیل پروژه و هزینه کاهش یابد. زمان شروع و پایان هر وظیفه[1] با استفاده از رابطه ۱و۲ انجام می‌گیرد.

$$t_j^{start} = \begin{cases} 0 & if \forall k \neq j, (t_k, t_j) \notin A \\ \max\{t_k^{end} \middle| (t_k, t_j) \in A\} & else \end{cases} \quad (1)$$

$$t_j^{end} = t_j^{start} + t_j^{duration} \quad (2)$$

محاسبه مدت زمان اجرای هر task با استفاده از رابطه ۳ انجام می‌گیرد.

$$t_j^{duration} = \frac{t_j^{effort}}{\sum_{i=1}^{E} m_{ij}} \quad (3)$$

محاسبه مدت زمان اجرای پروژه با استفاده از رابطه ۴ انجام می‌گیرد.

$$p^{duration} = \max\{t_j^{end} \middle| \forall k \neq j (t_j, t_k) \notin A \quad (4)$$

محاسبه هزینه برای هر وظیفه با استفاده از رابطه ۵ انجام می‌گیرد.

$$t_j^{\cos t} = \sum_{i=1}^{E} e_i^{salary} . m_{ij} . t_j^{duration} \quad (5)$$

محاسبه هزینه تمام وظیفه های پروژه با استفاده از رابطه ۶ انجام می‌گیرد.

$$p^{\cos t} = \sum_{j=1}^{T} t_j^{\cos t} \quad (6)$$

کارآیی کارکنان با استفاده از روابط ۸،۷ و۹ انجام می‌گیرد.

$$e_i^{over} = \int_{t=0}^{t=p^{duration}} ramp(e_i^{work}(t) - e_i^{\max ded}) dt \quad (7)$$

$$where, \qquad ramp(x) = \begin{cases} x & if x > 0 \\ 0 & if x \leq 0 \end{cases} \quad (8)$$

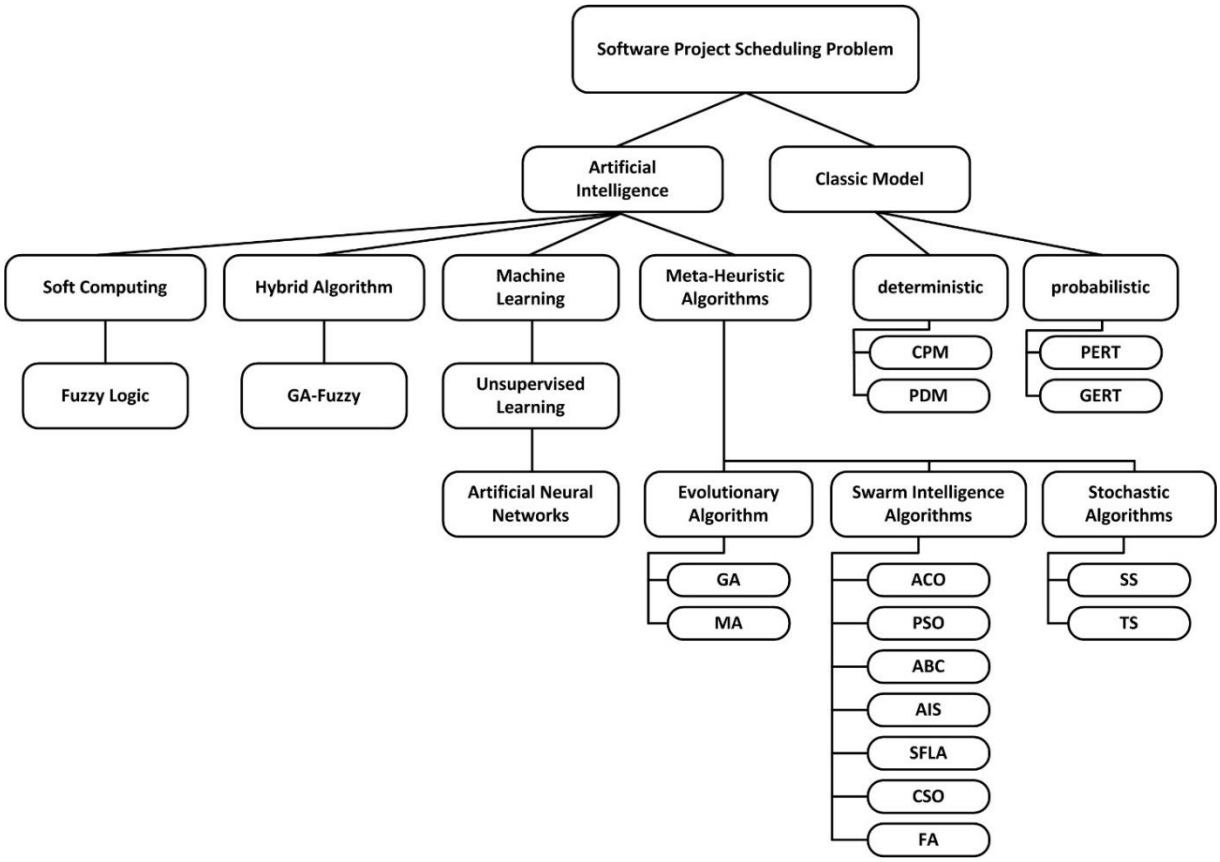
$$e_i^{work}(t) = \sum_{\{j \middle| t_j^{start} \leq t \leq t_j^{end}} m_{ij} \quad (9)$$

محاسبه کار کلی برای پروژه‌های نرم‌افزار با استفاده از رابطه ۱۰ انجام می‌گیرد.

$$p^{over} = \sum_{i=1}^{E} e_i^{over} \quad (10)$$

منابع انسانی و بودجه باید بصورت بهینه مدیریت شوند تا پروژه با موفقیت پایان پذیرد. در چند سال اخیر تحقیقات زیادی، در زمینه زمانبندی پروژه‌های نرم‌افزاری با استفاده از تکنیک‌های هوش مصنوعی صورت گرفته است. اما هنوز با قطعیت کامل نمی‌توان گفت که روش‌های هوش مصنوعی ۱۰۰٪ مشکلات هزینه و زمان را پروژه‌های نرم‌افزاری برطرف خواهند کرد. اما طبق بررسی‌هایی که ما انجام دادیم، به این نتیجه رسیدیم که تکنیک‌های هوش مصنوعی کارایی بهتری در مقابل تکنیک‌های کلاسیک زمانبندی از خود نشان داده‌اند. در شکل۱ دسته‌بندی الگوریتم‌هایی که به حل زمانبندی پروژه‌های نرم‌افزاری پرداخته‌اند نشان داده شده است.

[1] Task



شکل۱: دسته‌بندی مدل‌های مختلف در SPSP

Figure 1. Classification of different models in SPSP

۲-۱- مدل‌های کلاسیک

مدل PERT: این روش در مواردی که اطلاعاتی از تجربیات گذشـــته برای تخمین پروژه وجود ندارد، به‌کار می‌رود [۷]. از این روش در برنامه‌ریزی و کنترل تولید محصولات جدید و پروژه‌های تحقیقاتی که عوامل نامشخص در خصوص هزینه و زمانبندی باید ارزیابی گردند، استفاده می‌شود. مدل PERT بر این اصل متکی است که باید پروژه‌های نرم‌افزاری در زمان‌های تعیین شده به اتمام برسند. سیستم تخمین برای مدت زمان وظایف در مدل PERT به صورت رابطه ۱۱ است.

$$T = \frac{(T_o + 4T_m + T_p)}{6} \qquad (۱۱)$$

To: عبارت است از مطلوب‌ترین زمان[1] که یک فعالیت در شرایط یکسان صرف خواهد نمود.

Tm: عبارت است از محتمل‌ترین زمان[2]که از حداکثر فراوانی در تابع توزیع زمان برخوردار است.

Tp: عبارت است از نامطلوب‌ترین زمان[3] که یک فعالیت در شرایط یکسان صرف خواهد نمود.

بنابراین، با اســتفاده از مدل PERT، تحلیلگر قادر اســت، مطلوب‌ترین مدت را برای پروژه، احتمال زمان ترســیم کل پروژه و یا بخشی از آن را تخمین نماید.

مدل GERT: در صورتی که در روش GERT، ترکیب انجام وظایف احتمالی بوده و فاقد قاطعیت است [۸].

مدل CPM: در مدل CPM مدت زمان مورد نیاز برای انجام فعالیت‌ها قطعی فرض می‌گردد [۹]. مدل CPM کل فعالیت‌های پروژه و روابط میان آنها را نشـــان می‌دهد. مدیر پروژه، تحلیلگران، برنامه‌نویســـان و افراد ســـهیم در اجرای پروژه، با در اختیار

[1] Optimistic Time
[2] Most Likely Time
[3] Pessimistic Time

داشتن شبکه، می‌توانند تصورات و فرضیات قبلی خود درباره فعالیت‌های پروژه و روابط میان آنها را به طور عینی بررسی کرده و آنها را تغییر دهند.با مدل‌های PERT و CPM، می‌توان مسائلی نظیر تخصیص منابع و کاهش هزینه و نیروی انسانی را مورد تجزیه و تحلیل قرار داد. روش PERT در خصوص آن دسته از وظایف کاربرد دارد که مدت زمان لازم برای انجام آنها قطعی ناست.پس از عنوان مقاله باید نام نویسندگان نوشته شوند. از ذکر عناوینی مثل استاد، دکتر، مهندس، و غیره خودداری کنید. نام دانشگاه یا محل اشتغال نویسنده به همراه نشانی و آدرس پست الکترونیکی می‌توانند ذکر شوند. نام نویسندگان به زبان انگلیسی نیز باید نوشته شود.

مدلPDM [1]: مدل PDMبه منظور برطرف کردن مشکلات مدل CPM ابداع گردید [۹]. در این مدل می‌توان شبکه را با کمترین تعداد ارتباط بین وظایف طراحی کرد. مدل PDM توانایی بیشتری برای گسترده کردن پروژه‌ها دارد و براحتی بروزرسانی می‌شود. در جدول۲ مزایا و معایب مدل‌های کلاسیک بحث شده است.

جدول۲: مزایا و معایب مدل‌های کلاسیک

Table 2. Advantages and disadvantages of classic models

مدل‌های کلاسیک				
توضیحات	**معایب**	**مزایا**	**ایده اصلی**	**رویکرد**
با استفاده از مدل PERT، می‌توان محتمل‌ترین مدت زمان پروژه و احتمال زمان ترسیم کل پروژه و یا بخشی از آن را برآورد کرد.	۱. عدم تشخیص دقیق روابط بین وظایف ۲. عدم امکان تعیین اثرات تاخیرات در پروژه‌های نرم‌افزاری	۱. شناسایی منابع مورد نیاز در ابتدای اجرای پروژه‌ها ۲. تجزیه و تحلیل هزینه برای هر یک از وظایف. ۳. تجزیه و تحلیل هزینه مرتبط با تأخیرها و تغییرات ۴. مشخص شدن میزان کارآیی کارکنان و سرعت پیشرفت پروژه	در این روش زمان انجام وظایف احتمالی است.	PERT
روش GERT برای پروژه‌هایی بکار می‌رود که اجرا یا عدم اجرای آنها به شکل احتمالی است و در شروع پروژه قابل پیش‌بینی نیستند.	۱. تحلیل زمانبندی دشوار است. ۲. افزایش هزینه	۱. ترسیم روابط بین وظایف و شناسایی رابطه میان وظایف ۲. محاسبه تخمین زمان برای هرکدام از وظایف	زمان انجام وظایف و وقوع وظایف احتمالی است	GERT
در مدل CPM با استفاده و توجه کامل به مدت زمان، ارتباطات، وابستگی‌ها و توالی فعالیت‌ها، زودترین و دیرترین زمان شروع و خاتمه هر فعالیت، به طور قطعی، تعیین و مشخص می‌شود.	۱. زمانبر بودن ۲. افزایش پیچیدگی روش با افزایش تعداد وظایف	۱. برآورد زمان و هزینه برای هر کدام از وظایف ۲. مشخص کردن مراحل اصلی و بحرانی و نقاط تحویل در پروژه‌ها ۳. دستیابی به نتایج قابل اطمینان	در این روش زمان انجام وظایف قطعی است.	CPM
در مدل PDM زمان، توالی و ارتباط فعالیت‌ها، شروع و پایان هر فعالیت نشان داده می‌شود.	۱. به دلیل وجود فرمولهای بیشتر نسبت به روش CPM پیچیده‌تر است. ۲. افزایش پیچیدگی	۱. ارتباطات کمتری در مقایسه با CPM بین وظایف وجود دارد. ۲. در مدل CPM هیچگونه روابط Dummy بین وظایف وجود ندارد. ۳. امکان تعریف روابط جدید بین فعالیت‌ها	در این روش زمان انجام وظایف قطعی است.	PDM

[1] Precedence Diagram Method

۲- ۲- الگوریتم‌های فرااکتشافی

تلاش‌ها و تحقیقات صورت گرفته توسط محققین در سالهای اخیر منجر به ابداع الگوریتم‌هایی با الهام از پدیده‌های طبیعت شد، پدیده‌هایی که به بررسی تکامل و رفتار موجودات طبیعت می‌پردازد و منجر به الگوریتم‌های فرااکتشافی شدند. الگوریتم‌های فرااکتشافی که بر مبنای جمعیت به حل مسائل بهینه‌سازی می‌پردازند، ابزارهایی برای یافتن راه حلهای نزدیک به بهینه هستند. این الگوریتم‌ها با بهره بردن از دو مفهوم تنوع[1] و همکاری، فضای مسائل بهینه‌سازی را تا رسیدن به بهینه‌ترین حالت، جستجو می‌کنند. به دلیل افزایش پیچیدگی مسائل بهینه‌سازی و ناتوانی مدل‌های الگوریتمی برای ارائه راه حل بهینه، الگوریتم‌های فرااکتشافی راهکار مناسبی برای مسائل بهینه‌سازی هستند. مدل‌های الگوریتمی در حل بسیاری از مسائل علمی و مهندسی مورد استفاده قرار گرفته‌اند و دامنه بسیار وسیعی از مسائل مختلف را پوشش می‌دهند اما مدل‌های الگوریتمی با وجود کارایی دقیق، هنوز هم در حل بسیاری از مسائل بهینه‌سازی با مشکلات زیادی مواجه هستند.

۲-۲-۱- الگوریتم ژنتیک

الگوریتم ژنتیک[2] اولین بار در سال ۱۹۷۵ توسط جان هلند معرفی شد [۱۰]. الگوریتم ژنتیک، با در نظر گرفتن مجموعه‌ای از نقاط فضای جواب در هر تکرار محاسباتی، به نحو مؤثری، نواحی مختلف فضای جواب را جستجو می‌کند. این الگوریتم با ایجاد جمعیت اولیه به حل مسائل می‌پردازد. هر یک از افراد[3] جمعیت که کروموزوم[4] نامیده می‌شوند یک جواب برای مساله مورد نظر محسوب می‌شود و هر یک از اجزای کروموزوم‌ها که ژن نامیده می‌شوند بیانگر متغیر خاصی از مساله بهینه‌سازی مورد نظر هستند. نسل جدید با در نظر گرفتن تابع برازش افراد و با به کارگیری عملگرهای الگوریتم ژنتیک (ادغام و جهش) تولید می‌شود و تابع برازش افراد در طول تکرار الگوریتم بهبود می‌یابد.

در [۱۱] یک الگوریتم ژنتیک برای مسئله زمانبندی پروژه نرم‌افزاری با محدودیت منابع[5] ارائه شده‌است. الگوریتم پیشنهادی چندین تغییر را در الگوی الگوریتم ژنتیک، مانند یک اپراتور انتخاب جدید برای انتخاب وللدین، یک عملگر تقاطع دو نقطه‌ای اصلاح‌شده با مرتبه تقاطع خاص و یک عملگر جهش مبتنی بر احتمال کاهش خطی معرفی می‌کند. الگوریتم پیشنهادی با استفاده از مسایل معیار استاندارد اندازه J30، J60 و J120 از کتابخانه مساله زمانبندی پروژه مورد آزمون قرار گرفته و مقایسه شده است. در [۱۲] مسأله زمانبندی پروژه چند هدفه ژنتیک با محدودیت منابع بررسی شده است. به منظور نزدیکی هر چه بیشتر مدل به شرایط واقعی زمان انجام هر فعالیت به صورت چند هدفه در نظر گرفته شده است. اهدافی که در این مدل در نظر گرفته شده‌اند عبارتند از حداقل کردن زمان و هزینه کل پروژه. با توجه به چند هدفه بودن و پیچیدگی محاسباتی مدل بدست آمده، از الگوریتم تکاملی چند هدفه معروف بنام الگوریتم ژنتیک[6] NSGA-II استفاده شده است. برای ارزیابی روش پیشنهادی، ۳۶ مساله زمانبندی پروژه انتخاب شده و کارایی NSGA-II بر پایه شاخص‌های طراحی شده، با الگوریتم‌های PAES[7] و SPEA2[8] مورد مقایسه قرار گرفته شده است. و در نهایت نتایج بررسی‌ها نشان می‌دهد الگوریتم NSGA-II کاراتر است.

الگوریتم ژنتیک فوق اکتشافی[9] برای مسئله زمانبندی پروژه در [۱۳]پیشنهاد شده است. این پژوهش در طی سه مرحله به حل این مسئله می‌پردازد. در مرحله اول، یک بردار توالی کار برای رمزگذاری راه حل استفاده می‌شود و یک طرح رمزگشایی مبتنی بر تعمیر برای تولید زمان بندی های امکان پذیر پیشنهاد شده است. در مرحله بعدی ده قانون اکتشافی ساده برای ساخت مجموعه ای از اکتشافی‌های سطح پایین طراحی شده‌اند. سپس، برنامه‌ریزی ژنتیکی به‌عنوان یک استراتژی سطح بالا استفاده می‌شود که می‌تواند اکتشافی‌های سطح پایین در حوزه اکتشافی را به طور انعطاف‌پذیر مدیریت کند. علاوه بر این، از روش

---

[1] Diversity
[2] Genetic Algorithm
[3] Individuals
[4] Chromosome
[5] Resource Constrained Project Scheduling Problem (RCPSP)
[6] Non-dominated Sorting Genetic Algorithm
[7] Pareto Archived Evolution Strategy
[8] Strength Pareto Evolutionary Algorithm 2
[9] Genetic Programming Hyper-Heuristic (GP-HH)

طراحی آزمایش[1] برای بررسی تأثیر تنظیم پارامترها استفاده می‌شود. در نهایت، عملکرد GP-HH بر روی مجموعه داده‌های معیار چند هدفه پروژه متشکل از ۳۶ مورد ارزیابی می‌شود. مقایسه محاسباتی بین GP-HH و الگوریتم های پیشرفته نشان دهنده برتری GP-HH پیشنهادی در محاسبه راه حل های امکان پذیر برای مشکل است.

۲-۲-۲- الگوریتم ممتیک

الگوریتم ممتیک[2] برای رفع مشکلات الگوریتم ژنتیک ارائه گردید [۱۴]. الگوریتم ژنتیک در حل بسیاری از مسائل می‌تواند کارایی خوبی داشته باشد. ولی در مواقعی که فضای جستجو بسیار بزرگ است، سرعت همگرائی این الگوریتم برای رسیدن به راه حل بهینه بسیار آهسته است. الگوریتم ممتیک یک الگوریتم تکاملی با جستجوی محلی می باشد. الگوریتم ممتیک در واقع حاصل ترکیب الگوریتم ژنتیک و جستجوی محلی می باشد. به عبارتی در این الگوریتم جستجوی محلی برای هر فرزند تولید شده در الگوریتم ژنتیک استفاده می‌شود تا فرزندان را به سوی جواب بهینه محلی هدایت نماید. محققین در مرجع [۱۵] رویکردی جدید بر مبنای الگوریتم ممتیک به نام الگوریتم ممتیک بدون مقیاس[3] برای RCPSP ارائه داده‌اند. الگوریتم پیشنهادی بر روی دیتاست‌های j30 وj60 مورد تست و آزمایش قرار گرفته است. نتایج آزمایشات نشان می‌دهند که مدل SFMS عملکرد بالایی دارد.

یک الگوریتم ممتیک مبتنی بر الگوریتم ژنتیک برای حل RCPSP در [۱۶] ارائه شده است. این الگوریتم شامل عملیات تقاطع کلاسیک و یک جستجوی محلی مبتنی بر درج متغیر است. یک طرح راه اندازی مجدد خودکار نیز ارائه شده است که به الگوریتم کمک می‌کند تا از بهینه محلی خارج گردد. بعلاوه، یک روش طراحی آزمایش برای تعیین مجموعه پارامترهای مناسب برای الگوریتم ممتیک پیشنهادی استفاده می‌شود. نتایج عددی، تجزیه و تحلیل آماری و مقایسه با الگوریتم‌های پیشرفته، اثربخشی رویکرد پیشنهادی را نشان می‌دهد.

۲-۲-۳- بهینه‌سازی کلونی مورچگان

الگوریتم کلونی مورچگان[4] از شناخته شده ترین الگوریتم‌های فرااکتشافی است. الگوریتم کلونی مورچگان در سال ۱۹۹۶ ارائه شد [۱۷]. این الگوریتم برگرفته از زندگی مورچه‌های طبیعی است. مورچه‌ها هنگام راه رفتن در مسیر خود ماده بوداری به نام فرومون[5] بجای می‌گذارند. البته این ماده بزودی تبخیر می‌شود ولی در کوتاه مدت به عنوان رد مورچه بر سطح زمین باقی می‌ملند. مورچه‌ها این قابلیت را دارند که می‌توانند با تولید فرومون، کوتاه‌ترین مسیر به غذا را پیدا کنند. مورچه‌هایی که کوتاه‌ترین مسیر را انتخاب می‌کنند، نسبت به آن‌هایی که مسیر طولانی‌تری را انتخاب می‌کنند، فرومون بیشتری، ایجاد می‌کنند. از آنجاکه فرومون بیشتر، مورچه‌ها را بهتر جذب می‌کند، مورچه‌های بیشتر و بیشتری، مسیر کوتاه‌تر را انتخاب می‌کنند تا آنجاکه همه مورچه‌ها، کوتاهترین مسیر را یافته و از آن مسیر حرکت می‌کنند. برای بررسی بیشتر موضوع، فرض می‌کنیم که به عنوان مثال، دو مسیر به منبع غذا وجود دارند که دارای طول متفاوتی هستند. مورچه‌ها، هر دو مسیر را با احتمالات یکسان، انتخاب می‌کنند. مورچه‌هایی که مسیر کوتاه‌تر را رفته و برگشته‌اند، بیشترین فرومون را زودتر از بقیه تولید می‌کنند. در نتیجه، مورچه‌های دیگر این مسیر را زودتر انتخاب کرده و فرومون این مسیر را بیشتر تقویت می‌کنند. در نهایت همه مورچه‌ها، کوتاه‌ترین مسیر به غذا را می‌پیمایند.

در [۱۸] از الگوریتم بهینه‌سازی کلونی مورچه‌ها که از الگوریتم‌های فرااکتشافی است، برای حل SPSP استفاده شده است. ارزیابی از جهات مختلفی مانند زمان شروع و پایان هر پروژه، زمان کلی پروژه و هزینه کلی پروژه انجام گرفته است. برای اینکه کارایی الگوریتم کلونی مورچگان به خوبی نشان داده شود از الگوریتم ژنتیک برای مقایسه کارایی استفاده شده است. در بخش نتایج، فاکتورهای نحوه بروزرسانی فرومون بر روند بهینه‌سازی الگوریتم، تعداد کارکنان، تعداد وظایف و مهارت‌های کارکنان در نظر گرفته شده است. نتایج آزمایشات نشان می‌دهد که کارایی الگوریتم کلونی مورچگان در تخمین هزینه و زمان در مقایسه با

[1] Design of Experiments
[2] Memetic Algorithm
[3] Scale-Free Based Memetic Algorithm (SFMS)
[4] Ant Colony Optimization
[5] Pheromone

الگوریتم ژنتیک بهتر است. برخی از محققین [۱۹] برای زمانبندی پروژه‌های نرم‌افزاری از الگوریتم کلونی مورچگان استفاده کرده‌اند.

یکی از مسائل بسیار مهم در مدیریت پروژه‌های نرم‌افزاری، انتخاب بهترین راه حل برای انجام هر کدام از فعالیت‌های تشکیل دهنده پروژه است، به نحوی که هزینه و زمان پایانی پروژه دارای کمترین مقدار ممکن باشد. با توجه به تعداد زیاد فعالیتها و راه حل‌های انتخابی برای هر فعالیت، معمولاً این انتخاب دارای یک جواب منحصر به فرد نیست بلکه مجموعه‌ای از جوابها را تشکیل می‌دهد، که این مجموعه جواب‌ها هیچکدام بر دیگری ترجیحی ندارند. از سوی دیگر در پروژه‌های واقعی معمولاً هزینه‌های پیش‌بینی شده برای انجام فعالیتها همراه با عدم قطعیت‌هایی هستند که منجر به تغییرات زیادی در هزینه تمام شده پروژه می‌شوند. در روش‌های کلاسیک جهت زمانبندی پروژه‌ها، کاهش زمان پروژه نسبت به هزینه در اولویت بالاتری قرار می‌گیرد، در حالی‌که در بسیاری موارد کاهش زمان پروژه سبب افزایش هزینه می‌گردد. آنها با ارائه الگوریتم کلونی مورچگان روند زمانبندی، فضای جستجوی زمانبندی را گسترده‌تر کرده‌اند و همچنین برای اینکه کارایی الگوریتم کلونی مورچگان به خوبی نشان داده شود نتایج آن را با الگوریتم ژنتیک مقایسه کرده‌اند و در بسیاری موارد الگوریتم کلونی مورچگان حالت بهینه یا نزدیک به بهینه را پیدا می‌کند.

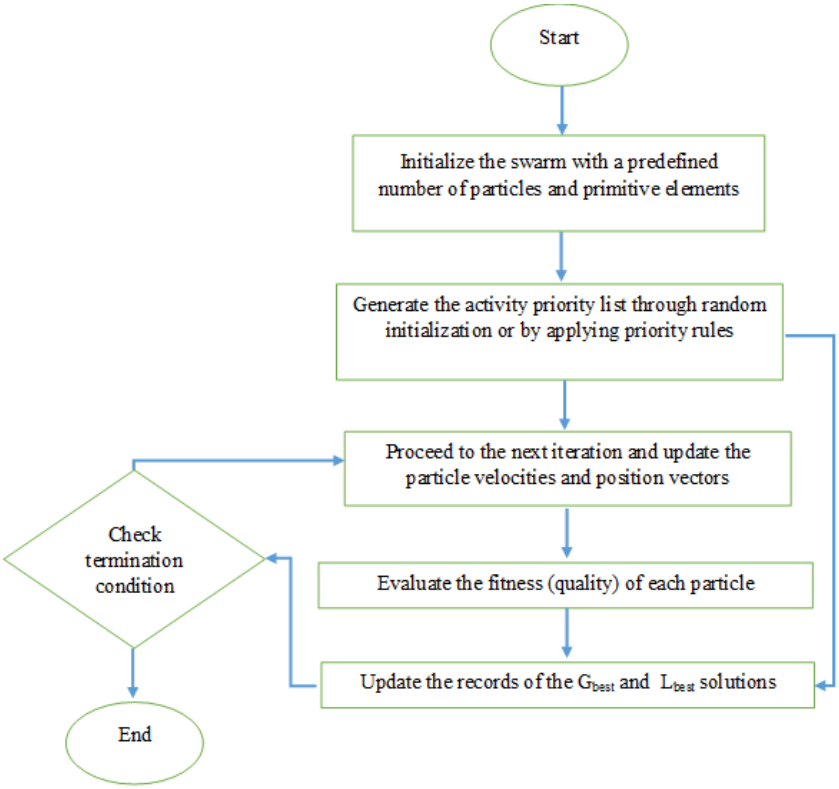
در الگوریتم پیشنهادی[۲۰] مسئله بهینه‌سازی به چندین مشکل فرعی و مورچه‌های موجود در جمعیت به مورچه‌های نخبه و مورچه‌های معمولی تقسیم می‌شوند تا نرخ همگرایی بهبود یابد و از قرار گرفتن در مقدار بهینه محلی جلوگیری شود. استراتژی به‌روزرسانی فرمون برای بهبود توانایی بهینه سازی استفاده می‌شود. مکانیسم انتشار فرمون برای ساختن فرمون آزاد شده توسط مورچه‌ها در یک نقطه خاص استفاده می‌شود که به تدریج محدوده خاصی از مناطق مجاور را تحت تاثیر قرار می‌دهد. مکانیسم تکامل مشترک برای تبادل اطلاعات بین جمعیت‌های فرعی مختلف به منظور اجرای اشتراک‌گذاری اطلاعات استفاده می‌شود. به منظور تأیید عملکرد بهینه‌سازی الگوریتم مورد بحث، مسئله فروشندگان دوره‌گرد و مسئله تخصیص گیت واقعی در اینجا انتخاب شده‌اند. نتایج آزمایش نشان می‌دهد که الگوریتم پیشنهادی می‌تواند به طور موثر بهترین مقدار بهینه‌سازی را در حل فروشندگان دوره‌گرد به دست آورد و به طور موثر مسئله انتساب گیت را حل کند، نتیجه تخصیص بهتر را به دست آورد و توانایی بهینه‌سازی و پایداری بهتری را به دست آورد.

۲-۲-۴- بهینه‌سازی اجتماع ذرات

الگوریتم بهینه‌سازی ازدحام ذرات[1] در سال ۱۹۹۵ با الهام از رفتار اجتماعی پرندگان که در گروه‌های کوچک و بزرگ زندگی می‌کنند، معرفی شد [۲۱]. بهینه‌سازی ازدحام ذرات، یک شبیه‌سازی از رفتار اجتماعی گروهی از پرندگان است، که در محیطی به دنبال غذا می‌گردند. هیچ یک از پرندگان در مورد محل غذا اطلاعاتی ندارند ولی در هر مرحله می‌دانند که چقدر از محل غذا فاصله دارند. بر این اساس، بهترین رویکرد برای پیدا کردن غذا، پیروی از نزدیکترین پرنده به غذا است. الگوریتم بهینه‌سازی ازدحام ذرات یک الگوریتم جمعیتی است که در آن تعدادی ذره که راه حل‌های یک تابع یا یک مسئله هستند، یک ازدحام (جمعیت) را تشکیل می‌دهند. جمعیتی از ذرات در فضای مسئله حرکت کرده و براساس تجربیات فردی خود و تجربیات جمعی سعی می‌کنند تا راه حل بهینه در فضای جستجو را بیابند. الگوریتم بهینه‌سازی ازدحام ذرات به عنوان یک الگوریتم بهینه‌سازی، یک جستجوی مبتنی بر جمعیت را فراهم می‌کند که در آن هر ذره با گذشت زمان موقعیت خود را تغییر می‌دهد. در الگوریتم بهینه‌سازی ازدحام ذرات، ذرات در یک فضای جستجوی چند بعدی از راه‌حل‌های ممکن مساله، حرکت می‌کنند. در این فضا یک معیار ارزیابی تعریف می‌شود و سنجش کیفیت راه‌حل‌های مساله از طریق آن صورت می‌پذیرد. تغییر حالت هر ذره در یک گروه تحت تاثیر تجربه‌های خود و یا دانش همسایگانش بوده و رفتار جستجوی یک ذره در گروه تحت تاثیر ذرات دیگر است. همین رفتار ساده باعث پیدا شدن ناحیه‌های بهینه از فضای جستجو می‌گردد. بنابراین در الگوریتم بهینه‌سازی ازدحام ذرات هر ذره به محض پیدا کردن موقعیت بهینه آن را به نحو مناسبی به اطلاع سایر ذرات می‌رساند و هر ذره بر اساس مقادیر بدست آمده برای تابع هزینه با یک احتمال معین تصمیم می‌گیرد تا از ذرات دیگر پیروی نماید و جستجو

[1] Particle Swarm Optimization

در فضای مساله با استفاده از دانش قبلی ذرات انجام گیرد. این عمل باعث می‌شود تمام ذرات بیش از حد به یکدیگر نزدیک نشوند و به طور موثری از عهده حل مسائل بهینه‌سازی برآیند. شکل ۲ نمودار استفاده از الگوریتم ازدحام ذرات را برای حل مساله زمانبندی را نشان می‌دهد.



شکل ۲: فلوچارت عمومی بهینه‌سازی اجتماع ذرات برای مساله زمان بندی[۲۲]

Figure 2. General flowchart of particle swarm optimization for scheduling problem [22]

مرجع [۲۳] محققین از الگوریتم بهبود یافته بهینه‌سازی ازدحام ذرات برای حل RCPSP استفاده کرده اند. با تغییر در نحوه بروزرسانی و موقعیت ذرات سعی شده است نقاط بهتری از فضای جستجو پیدا شوند. الگوریتم بهینه‌سازی ازدحام ذرات بهبود یافته بر روی داده‌های j30، j60، j90و j120مورد تست و ارزیابی قرار گرفته است. نتایج آزمایشات نشان می‌دهد که مدل بهینه‌سازی ازدحام ذرات بهبود یافته در مقایسه با الگوریتم بهینه‌سازی ازدحام ذرات توانسته است مقدار انحراف میانگین را از ۱.۳۳ به ۰.۴۹ کاهش دهد. در [۲۴]، یک الگوریتم فراابتکاری بر اساس الگوریتم بهینه‌سازی ازدحام ذرات جهت تولید جوابهای مناسب برای مساله زمان‌بندی RCPSP با در نظر گرفتن محدودیت منابع و فعالیتهایی با زمان‌های اجرای قطعی و احتمالی توسعه داده شده است. این الگوریتم از ذرات و دانش ذرات و جواب‌های اولیه مختلف بعنوان راهکاری برای ایجاد تنوع در فضای جستجو استفاده می‌نماید. در طی مثال‌های مختلف مشخص می‌شود که الگوریتم بهینه‌سازی ازدحام ذرات بعنوان الگوریتمی کارا در ایجاد جواب‌های مناسب برای زمان‌های قطعی و احتمالی کاربرد دارد. الگوریتم بهینه‌سازی ازدحام ذرات بر روی داده‌های j30، j60، j90و j120مورد تست قرار گرفته است. نتایج آزمایشات نشان می‌دهد که الگوریتم بهینه‌سازی ازدحام ذرات در مقایسه با الگوریتم جستجوی ممنوعه ,الگوریتم ژنتیک و شبیه سازی حرارت دقت بهتری در زمانبندی وظایف دارد. مدیران پروژه همواره با چالش تخصیص بهینه منابع به پروژه‌ها مواجه بوده‌اند تا بدین وسیله بین اهداف مختلف و اغلب در تضاد پروژها توازن برقرار سازند. زمان، هزینه و کیفیت تحویل پروژه‌ها جزء اهداف مهم هر پروژه‌ای می‌باشند. در مرجع [۲۵] یک روش حل مبتنی بر الگوریتم بهینه‌سازی ازدحام ذرات برای RCPSP پیشنهاد شده است. در این مرجع به منظور کارایی بهتر، الگوریتم به PSO+ بهبود داده شده است. ارزیابی و نتایج بر روی داده‌های j30، j60، j90و j120از دیتاست PSPLIB انجام گرفته است. نتایج آزمایشات نشان می‌دهد که الگوریتم PSO+ کارایی بهتری در مقایسه با الگوریتم‌های بهینه‌سازی

ازدحام ذرات‌و کلونی مورچگان دارد. منطق فازی با در نظر گرفتن عدم قطعیت در پارامترهای تصمیم‌گیری و استفاده از مدل‌های انتزاعی، رویکردی فازی جهت نزدیک کردن مدل‌های زمانبندی پروژه‌های نرم‌افزاری به واقعیت است.

در [۲۶] یک الگوریتم بهینه‌سازی ازدحام ذرات مبتنی بر الگوریتم فوق ابتکاری برای حل مساله زمانبندی پروژه با محدودیت منابع پیشنهاد شده‌است. روش فوق ابتکاری به عنوان یک الگوریتم سطح بالا عمل می‌کند که چندین روش ابتکاری سطح پایین را کنترل می‌کند که در فضای راه‌حل عمل می‌نمایند. نمایش راه‌حل براساس کلیدهای تصادفی است. برنامه‌های زمان‌بندی فعال با استفاده از برنامه زمان‌بندی سریال با استفاده از اولویت‌های فعالیت‌هایی که با ابتکارات سطح پایین الگوریتم اصلاح می‌شوند، ساخته می‌شوند. همچنین، عملگر توجیه دوگانه، یعنی یک روش بهبود رو به جلو، برای همه راه‌حل‌ها اعمال می‌شود. روش پیشنهادی بر روی مجموعه‌ای از نمونه‌های مساله استاندارد کتابخانه Wellknown مورد آزمایش قرار گرفت و با روش‌های دیگر مقایسه شد. نتایج محاسباتی نویدبخش کارایی روش پیشنهادی را تایید می‌کند.

در مقاله [۲۷] به مساله زمانبندی معکوس چند حالته کارگاه مونتاژ توربین در یک شرکت تولید تجهیزات نیروگاه برق چین با در نظر گرفتن برخی اختلالات غیر منتظره در فرآیند مونتاژ، مانند تاخیر مواد، شکست تجهیزات و غیره می‌پردازد. در روش پیشنهادی، کدگذاری دو بردار برای نشان دادن حالت‌های اجرایی فعالیت‌ها و زمان‌بندی اضافه‌کاری پروژه‌ها که به ترتیب توسط الگوریتم بهینه‌سازی ازدحام ذرات و جستجوی ممنوعه بهینه می‌شوند، ارائه شده است.یک الگوریتم کدگشایی اکتشافی ترکیبی[1] شامل قوانین انتخاب سفارش پروژه، قوانین زمانبندی جرثقیل، مکانیزم رزرو منابع و قواعد تعیین زمان اضافه پیشنهاد شده‌است. در نهایت با استفاده از داده‌های آزمایش تجربی و یک مورد مهندسی دنیای واقعی امکان پذیر بودن و موثر بودن الگوریتم پیشنهادی تایید می‌شود.

۲-۲-۵- بهینه‌سازی زنبور مصنوعی

الگوریتم بهینه‌سازی زنبور مصنوعی [۲][۲۸] یکی از الگوریتم‌های فرااکتشافی است که از رفتار کاوشی کلونی زنبورها برای حل مسائل بهینه‌سازی با فضاهای بزرگ الهام گرفته شده است. الگوریتم ABC با ایجاد جمعیت اولیه‌ای از بردارهای تصادفی شروع به کار می‌کند. نحوه کار بدین صورت است که در هر تکرار از الگوریتم، زنبورهای مصنوعی با انجام جستجوهای تصادفی در اطراف جوابهای بدست آمده در تکرار قبل به یافتن جوابهای جدید می‌پردازند. بدیهی است که این جوابهای جدید لزوما همگی بهتر از جوابهای بدست آمده در تکرار قبل نخواهد بود. پس از آنکه هر یک از زنبورهای مصنوعی مشغول به کار، یک جواب جدید پیدا کردند همه آنها دوباره به کندو باز می‌گردند و برای انتخاب مسیر حرکتشان در تکرار بعدی تصمیم‌گیری می‌کنند. بنابراین، میزان بهینگی جواب‌های بدست آمده توسط زنبورها محاسبه شده و سپس جوابی که برازندگی بیشتری دارد به عنوان مسیر جستجو در تکرار بعدی انتخاب می‌شود. بنابراین نواحی اطراف جوابهای بهینه‌تر توسط تعداد زنبورهای بیشتری در تکرار بعدی مورد جستجو قرار می‌گیرد. فرایند جستجو تا زمان تامین شرط مورد نیاز برای خاتمه اجرای برنامه انجام می‌گیرد.

در [۲۹] یک مدل ریاضی مساله و توصیف رفتار ABC ارائه شده‌است. همچنین مدلی را ارائه شده است که مساله را با ABC حل کند. از طریق این الگوریتم می‌توان راه‌حل‌هایی برای رقابت با سایر راه‌حل‌های ارائه‌شده توسط سایر نویسندگان ایجاد کرد.

در [۳۰] با استفاده از الگوریتم فرااکتشافی کلونی زنبورهای مصنوعی گسسته[3] به حل مساله RCPSP پرداخته شده است. از الگوریتم بهینه‌سازی زنبور عسل توزیع شده در جهت رسیدن به نتایج مطلوبتر نسبت به الگوریتم‌های ژنتیک، بهینه‌سازی ازدحام ذرات، شبیه‌سازی حرارت و جستجوی ممنوعه که تاکنون برای حل این مساله توسعه یافته‌اند تلاش شده است. الگوریتم بهینه‌سازی زنبور عسل توزیع شده بر روی داده‌های j30، j60، j90و j120 مورد تست قرار گرفته است. نتایج آزمایشات نشان می‌دهد که الگوریتم کرم شب‌تاب در مقایسه با الگوریتم‌های بهینه‌سازی ازدحام ذرات، جستجوی ممنوعه، الگوریتم ژنتیک و شبیه‌سازی حرارت دقت بهتری در زمانبندی وظایف دارد.

در [۳۱] سه الگوریتم زنبور عسل، کلونی زنبور عسل مصنوعی و بهینه‌سازی اجتماع زنبور از خانواده کلونی زنبورها برای مساله RCPSP ارائه شده است. از این الگوریتم‌ها در جهت رسیدن به نتایج مطلوبتر نسبت به الگوریتم‌های الگوریتم ژنتیک،

[1] hybrid heuristic decoding algorithm
[2] Artificial Bee Colony (ABC)
[3] Discrete Artificial Bee Colony

بهینه‌سازی ازدحام ذرات، شبیه‌سازی حرارت و جستجوی ممنوعه که تاکنون برای حل این مساله توسعه یافته‌اند تلاش شده است. ارزیابی و نتایج بر روی داده‌های j30، j60، j90 و j120 صورت گرفته است. نتایج آزمایشات نشان می‌دهد که الگوریتم‌های زنبور عسل در مقایسه با الگوریتم‌های بهینه‌سازی ازدحام ذرات، جستجوی ممنوعه، الگوریتم ژنتیک و شبیه‌سازی حرارت دقت بهتری در میانگین انحراف دارند. یکی از مسائل بهینه‌سازی ترکیبی RCPSP است. در [۳۲] به مسئله زمان‌بندی بدون انتظار توزیع ناهمگن با توجه به تفاوت انواع کارخانه‌ها، از جمله تعداد ماشین‌ها، تکنولوژی ماشین، تامین مواد خام و شرایط حمل و نقل پرداخته شده است تا زمان ساخت به حداقل برسد. در این مساله تعداد و نوع ماشین آلات در هر کارخانه متفاوت در نظر گرفته شده و این بدان معنی است که کارها باید از طریق مسیرهای پردازشی مختلف پردازش شوند. برای حل این مسئله، یک الگوریتم کلونی زنبورهای مصنوعی گسسته پیشنهاد شده است. در روش پیشنهادی کاهش همسایگی متغیر بر اساس چهار روش جستجوی محلی در فاز زنبور پیشاهنگ تعبیه شده تا توانایی جستجوی محلی الگوریتم کلی را تقویت کند.

۲-۲-۶- سیستم ایمنی مصنوعی

الگوریتم سیستم ایمنی مصنوعی[1] از سیستم ایمنی بدن موجودات زنده الهام گرفته شده است [۳۳]. کارایی سیستم ایمنی بدن موجودات زنده در تشخیص آنتی‌ژنها و از بین بردن آنها با تولید آنتی‌بادیهای لازم و مؤثر سبب شد تا محققان با شبیه‌سازی نحوه عملکرد این سیستم از توانایی آن در حل مسایل پیچیده مهندسی بهره گیرند. هنگامی که آنتی‌ژنها، ارگانیسم بدن را مورد تهاجم قرار می‌دهند، ضمن تخریب سلولها به تکثیر نیز می‌پردازند. بنابراین یکی از مکانیسم‌های جالب توجه سیستم تدافعی بدن موجودات زنده در مقابله با این تهاجم تکثیر سریع سلول‌های تدافعی است که توانایی شناسایی آنتی ژنها و نابودکردن آنها را دارا هستند. تشخیص یک آنتی‌ژن توسط آنتی‌بادیها با استفاده از فاکتور وابستگی[2] انجام می‌گیرد. آنتی‌بادیهایی که وابستگی کمتری با یکدیگر دارند با استفاده از عملگر جهش[3] با تغییرات ساختاری وابستگی خود را با عوامل بیماری‌زا بیشتر کرده و عملکرد دفاعی خویش را بهبودتر می‌بخشند.

در [۳۴] با استفاده از الگوریتم ایمنی مصنوعی به حل مساله زمانبندی پروژه‌ها پرداخته شده است. الگوریتم ایمنی مصنوعی بر روی داده‌های j30، j60، j90 و j120 از دیتاست PSPLIB مورد تست قرار گرفته است. نتایج آزمایشات نشان می‌دهد که الگوریتم AIS کارایی بهتری در مقایسه با الگوریتم‌های جستجوی پراکنده، ،شبیه‌سازی حرارت ، الگوریتم جستجوی ممنوعه و الگوریتم ژنتیک دارد. در [۳۵] با استفاده از الگوریتم ژنتیک مساله SPSP مورد بررسی قرار گرفته است. آنچه در این مرجع مطرح شده است مربوط به حالتی است که فعالیت‌ها در یک زمان قطعی از پیش تعیین شده به اتمام برسند. اگر فعالیت‌ها در زمان معمولی به اتمام برساند، اتمام پروژه‌ها در زمان ممکن امکان پذیر نیست. لذا باید با صرف منابعی با کمیت و کیفیت بهتری، زمان هر یک از فعالیت‌ها که تاخیر در اجرای آنها موجب تاخیر در اجرای کل پروژه است را کاهش داد. الگوریتم ژنتیک زمان هر یک از فعالیتها را کاهش داده و موجب موازنه زمان و هزینه در کنترل پروژه است. نتایج آزمایشات نشان می‌دهد که میانگین انحراف پروژه‌ها با استفاده از الگوریتم ژنتیک کاهش چشمگیری داشته است.

۲-۲-۷- الگوریتم جستجوی جهش قورباغه

الگوریتم جستجوی جهش قورباغه[4] یک الگوریتم جستجوی فرااکتشافی است که مبتنی بر جمعیت اولیه است و از تکامل طبیعی گروهی قورباغه‌ها که به دنبال محل با بیشترین ذخیره غذایی در دسترس می‌گردند، الهام گرفته شده است [۳۶]. محققین [۳۷] از الگوریتم جهش قورباغه برای مساله RCPSP استفاده کرده‌اند. این الگوریتم بر روی داده‌های j30، j60، j90 و j120 مورد تست قرار گرفته است. نتایج آزمایشات نشان می‌دهد که الگوریتم SFL با لحاظ کردن عدم قطعیت‌های موجود در پروژه‌ها در مقایسه با دیگر الگوریتم‌ها دقت بهتری دارد. زمانی که تعدادی از فعالیتهای پروژه به شیوه‌های متفاوتی و طبیعتاً با زمان و هزینه متفاوتی انجام شوند موجب می‌شود که پروژه‌ها نیز با هزینه‌ها و زمانهای مختلفی قابل اجرا باشد.

---

[1] Artificial Immune System
[2] Affinity
[3] Mutation
[4] Shuffled Frog Leaping Algorithm (SFL)

در [۳۸] با تاکید بر جستجو برای حل مسئله بهینه‌سازی چند هدفه و برنامه‌ریزی پروژه نرم‌افزاری راه حل جدیدی با استفاده از از الگوریتم پرش قورباغه ارائه شده است. روش ارائه شده در چهار مرحله نمایش قورباغه، راه حل، تعریف تابع برازندگی، پیاده‌سازی الگوریتم پرش قورباغه و ارزیابی برای یک مساله زمانبندی پروژه نرم‌افزاری به صورت تصادفی انجام شده است. نتایج نشان می‌دهد که روش پیشنهادی میزان خطا و هزینه را به مقدار قابل قبولی کاهش داده است.

۲-۲-۸- بهینه‌سازی جمعی گربه

الگوریتم بهینه‌سازی جمعی گربه‌ها[1] در سال ۲۰۰۷ با الهام از زندگی اجتماعی گربه‌ها که در گروه‌های جمعی زندگی می‌کنند، ابداع گردید [۳۹]. این الگوریتم مشابه الگوریتم‌های ژنتیک، کلونی مورچگان و بهینه‌سازی ازدحام ذرات است. مرجع [۴۰] یک رهیافت فرااکتشافی جدید با عنوان الگوریتم بهینه‌سازی جمعی گربه برای بهینه‌سازی RCPSP از لحاظ زمان و هزینه برای پروژه‌های مورد استفاده قرار داده است. الگوریتم بهینه سازی جمعی گربه از رفتار اجتماعی گربه‌ها پیروی می‌کند. کارایی الگوریتم در مقایسه با الگوریتم‌های فرااکتشافی دیگر بر روی دیتاست PSPLIB تست شده است. ارزیابی و نتایج بر روی داده‌های j30 وj60 صورت گرفته است. نتایج آزمایشات نشان می‌دهد که به طور میانگین الگوریتم بهینه‌سازی جمعی گربه کارایی بهتری دارد.

۲-۲-۹- الگوریتم کرم شب تاب

الگوریتم کرم شب‌تاب[2] [۴۱] یکی از جدیدترین الگوریتم‌های فرااکتشافی مبتنی بر هوش جمعی است که در حل مسائل بهینه‌سازی کاربرد دارد. در الگوریتم کرم شب‌تاب ابتدا تعدادی کرم شب تاب مصنوعی به طور تصادفی در دامنه مساله توزیع می‌شوند و سپس هر کرم شب تاب از خود نوری ساطع می‌کند که شدت آن متناسب با میزان بهینگی نقطه‌ای است که آن کرم شب تاب در آن واقع شده است. سپس شدت نور هر کرم شب تاب مرتباً با شدت نور سایر کرم‌های شب تاب مقایسه شده و کرم شب تاب کم نورتر به سوی کرم‌های شب تاب پر نورتر جذب می‌شود. همچنین پر نورترین کرم شب تاب نیز با هدف یافتن جواب بهینه سراسری به طور تصادفی در دامنه مساله حرکت می‌کند. بنابراین، در الگوریتم کرم شب‌تاب، کرم‌های شب‌تاب از طریق نور به تبادل اطلاعات با یکدیگر می‌پردازند. ترکیب این عملیات دسته جمعی باعث می‌شود که گرایش کلی کرم‌های شب تاب به سوی نقاط بهینه‌تر باشد. در [۴۲] روش جدیدی بر مبنای الگوریتم کرم شب‌تاب برای RCPSP ارائه شده است. دراین الگوریتم هر کرم شب‌تاب در واقع یک جواب شدنی برای تابع هدف است و براساس یک سری از توالی‌های ممکن عملیات فعالیت‌ها نشان داده می‌شود. مطابق محدودیت‌های مساله عملیات‌های هر کار براساس اولویت انجام فعالیت‌ها انجام می‌گیرد و مساله با هدف مینیمم کردن زمان اتمام کارها حل شده و یک توالی مشخص برای جواب بدست آمده پیشنهاد می‌گردد. ارزیابی و نتایج بر روی دیتاست PSPLIB انجام گرفته است. الگوریتم کرم شب‌تاب بر روی داده‌های j30، j60، j90 و j120 مورد تست قرار گرفته است. نتایج آزمایشات نشان می‌دهد که الگوریتم الگوریتم کرم شب‌تاب در مقایسه با الگوریتم‌های بهینه‌سازی ازدحام ذرات، جستجوی ممنوعه، الگوریتم ژنتیک و شبیه‌سازی حرارت دقت بهتری در زمانبندی دارد.

۲-۲-۱۰- جستجوی ممنوعه

الگوریتم جستجوی ممنوعه[3] یک الگوریتم بهینه‌سازی فرااکتشافی است که در سال ۱۹۸۶ معرفی شد [۴۳]. برای رسیدن به جواب بهینه در یک مسئله بهینه‌سازی، الگوریتم جستجوی ممنوعه ابتدا از یک جواب اولیه شروع به حرکت می‌کند. سپس الگوریتم بهترین جواب همسایه را از میان همسایه‌های جواب فعلی انتخاب می‌کند. در صورتی که این جواب در فهرست ممنوعه قرار نداشته باشد، الگوریتم به جواب همسایه حرکت می‌کند؛ در غیر این صورت الگوریتم معیاری به نام معیار تنفس را چک خواهد کرد. بر اساس معیار تنفس اگر جواب همسایه از بهترین جواب یافت شده تاکنون بهتر باشد، الگوریتم به آن حرکت خواهد کرد، حتی اگر آن جواب در فهرست ممنوعه باشد. پس از حرکت الگوریتم به جواب همسایه، فهرست ممنوعه بروزرسانی می‌شود؛ به این معنا که حرکت قبل که بوسیله آن به جواب همسایه حرکت کردیم در فهرست ممنوعه قرار داده می‌شود تا از

[1] Cat Swarm Optimization (CSO)
[2] Firefly Algorithm (FA)
[3] Tabu Search (TS)

بازگشت مجدد الگوریتم به آن جواب و ایجاد سیکل جلوگیری شود. در [۴۴] بر اساس الگوریتم جستجوی ممنوعه جهت تولید جواب‌های مناسب با در نظر گرفتن محدودیت منابع و حالت پیش نیازی و همچنین فعالیت‌هایی با زمانهای اجرای قطعی و احتمالی توسعه داده شده است. این الگوریتم از لیست‌های ممنوعه مختلف، حافظه تصادفی کوتاه مدت و جواب‌های اولیه مختلف به عنوان وسیله‌ای برای ایجاد تنوع در فضای جستجو استفاده می‌نماید. نتایج آزمایشات نشان می‌دهد که الگوریتم جستجوی ممنوعه به عنوان الگوریتمی کارا در ایجاد جواب‌های مناسب برای زمان های قطعی و احتمالی کارایی خوبی دارد.
در [۴۵] به مساله زمانبندی پروژه با محدودیت منابع پرداخته می‌شود که در آن منابع دارای مهارت‌های متعددی هستند و می‌توانند مهارت انتخاب‌شده را در لحظات زمانی گسسته تغییر دهند. هدف از این مقاله تعیین همزمان زمان شروع فعالیت‌ها و انتخاب مهارت‌های منابع برای بیشینه کردن استحکام زمانبندی و در عین حال ارضای محدودیت‌های اولویت، منابع تجدیدپذیر و مهلت پروژه است. پس از تعریف مساله، یک مدل برنامه‌ریزی خطی عدد صحیح ساخته شده و براساس ویژگی‌های مساله، یک الگوریتم تخصیص منابع برای تعیین مهارت منابع پیشنهاد شده است. دو الگوریتم خاص به عنوان دو معیار بهبود برای ترکیب با جستجوی تابو پایه در نظر گرفته می‌شوند که منجر به سه نسخه از جستجوی تابو می‌شود. با پیاده‌سازی روش پیشنهادی افزایش انعطاف‌پذیری منابع، قابلیت دسترسی منابع و مهلت پروژه و همچنین استحکام برنامه پایه افزایش می‌یابد.

۲-۲-۱۱- الگوریتم جستجوی پراکنده

الگوریتم جستجوی پراکنده[1] در قیاس با سایر الگوریتم‌های فرااکتشافی دارای ویژگی‌های متفاوتی است. این الگوریتم شامل مراحل تولید جمعیت اولیه، برازندگی، تشکیل زیرمجموعه و بروزرسانی مجموعه اولیه است [۴۶]. در [۴۷] راه حل‌های جدیدی برای مساله زمانبندی پروژه‌ها با استفاده از الگوریتم SS با هدف کمینه‌سازی حداکثر زمان تکمیل فعالیت‌ها و کاهش هزینه‌ها ارائه شده است. ارزیابی و نتایج بر روی داده‌های j30 وj60 صورت گرفته است. نتایج نشان می‌دهد که الگوریتم جستجوی پراکنده کارایی بسیار قابل قبولی در مقایسه با الگوریتم ژنتیک دارد. در [۴۸] با استفاده از الگوریتم جستجوی پراکنده به حل مساله زمانبندی پروژه‌ها پرداخته شده است. الگوریتم جستجوی پراکنده بر روی داده‌های j30، j60، j90 وj120از دیتاست PSPLIB مورد تست قرار گرفته است. نتایج آزمایشات نشان می‌دهد که الگوریتم جستجوی پراکنده کارایی خوبی برای مساله زمانبندی پروژه‌ها دارد. در [۴۹] بر مبنای ترکیب الگوریتم‌های جستجوی پراکنده و الکترومغناطیس یک مدل جدید برای مساله زمانبندی پروژه‌ها ارائه شده است. الگوریتم جستجوی پراکنده بر روی داده‌های j30، j60 وj120از دیتاست PSPLIB مورد تست قرار گرفته است. نتایج محاسباتی نشان می‌دهد که الگوریتم جستجوی پراکنده ترکیبی در کمینه‌سازی زمان کل پروژه‌ها کارایی اثربخشی دارد. جدول ۳ مقایسه الگوریتم‌های فرااکتشافی در زمانبندی پروژه‌های نرم‌افزاری را نشان می‌دهد.

۲-۳- الگوریتم‌های ترکیبی

الگوریتمی کاربردی و مناسب برای زمانبندی پروژه‌های نرم‌افزاری با استفاده از الگوریتم ژنتیک در شرایط فازی ارائه شده است [۵۰]. زمانبندی پروژه‌های نرم‌افزاری توسط متدولوژیهای سنتی همواره با عدم قطعیت مواجه هستند. الگوریتم ژنتیک با استفاده از ترکیب منطق فازی در بیان توالی مراحل تولید و عدم قطعیت فعالیت‌ها، الگوریتمی مناسب در زمانبندی پروژه‌های نرم‌افزاری است. در این مقاله پارامترهای زمان شروع هر فعالیت، زمان انجام هر فعالیت و تعداد تکرار حلقه ها در زمانبندی به شکل اعداد فازی مثلثی بیان می‌شوند. نتایج آزمایشات نشان می‌دهد ترکیب الگوریتم ژنتیک با منطق فازی نسبت به روش‌های موجود در زمانبندی پروژه‌های نرم‌افزاری کاربردی‌تر است و محاسبات کمتری دارد.
در [۵۱] با استفاده از منطق فازی روشی ترکیبی بر مبنای الگوریتم ژنتیک برای تخمین مدت زمان فعالیتها به صورت یک تابع فازی، ارائه شده است. برای زمانبندی پروژه‌های نرم‌افزاری باید به عدم اطمینان موجود در مورد بعضی پارامترهای پروژه‌ها مانند هزینه و زمان توجه شود. در این مرجع روشی برای زمانبندی پروژه‌ها با توجه به فازی بودن زمان فعالیت‌ها، محدودیت منابع و جستجو برای یافتن مناسب‌ترین برنامه زمانبندی فازی ارائه شده است.

[1] Scatter Search Algorithm (SS)

جدول۳: مقایسه الگوریتم‌های فرااکتشافی در زمانبندی پروژه‌های نرم‌افزاری

Table 3. Comparison of metahuristic algorithms in software project scheduling

الگوریتم‌های فرااکتشافی				
توضیحات	**معایب**	**مزایا**	**ایده اصلی**	**رویکرد**
الگورریتم ژنتیک با در نظر گرفتن مجموعه‌ای از نقاط فضای جواب در هر تکرار محاسباتی بطور موثری نواحی مختلف فضای جواب را جستجو می‌کند.	۱. گیرافتادن در بهینه سراسری ۲. پیچیدگی محاسباتی نسبتا زیاد است	۱. تنوع در اجرای وظایف ۲. دقت در برازندگی جوابهای بهینه ۳. همگرایی سریع	در این الگوریتم ، با توجه به میزان منابع موجود و زمان عادی وظایف، انواع برنامه‌های زمانبندی ممکن برای وظایف بررسی می‌گردند. بهینه‌ترین حالت در خروجی حاصل می‌گردد.	الگوریتم ژنتیک
این الگوریتم در مدت زمان کم، جواب‌های بهینه‌تری را نتیجه می-دهد.	۱. جستجوی محلی عموما از پیچیدگی بالایی برخوردار است. ۲. گیرافتادن در بهینه محلی	۱. احتمال گیرافتادن در نقاط بهینه محلی بسیار کم است. ۲. دامنه مساله به طور موثری جستجو می‌شود. ۳. سرعت همگرایی خوب ۴. کارایی و زمان اجرای این الگوریتم به مقدار پارامترهای آن وابسته است.	در این روش سعی می‌شود بهترین برنامه زمانبندی تهیه شود که هزینه و زمان آن مطلوب و قابل قبول باشد.	الگوریتم ممتیک
در الگوریتم بهینه‌سازی کلونی مورچه‌ها تمام فضای جواب به طور همه جانبه جستجو می‌شود، بنابراین امکان بیشتری برای همگرایی به بهینه محلی وجود خواهد داشت.	۱. گیر افتادن در بهینه محلی ۲. تعداد تکرار الگوریتم بری رسید به جواب بهینه زیاد است.	۱. پیدا کردن بهینه‌ترین حالت ۲. همگرایی الگوریتم ACO در مقایسه با الگوریتم ژنتیک بهتر است.	در این الگوریتم در صورتی که هیچ کدام از راه حل‌ها نتوانند حالت بهینه را پیدا کنند با تغییر زمان وظایف و میزان منابع موجود، برنامه‌های زمانبندی بهینه‌ترین بررسی می‌شوند.	بهینه-سازی کلونی مورچه
الگوریتم بهینه‌سازی اجتماع پرندگان برای گریز از بهینگی‌های محلی ترتیبی اتخاذ می‌کند که هنگام قرار گرفتن در بهینگی‌های محلی ذرات به بخش‌های دیگری از فضای جستجو جهش پیدا کنند و سپس در آنجا به جستجوی جواب بهینه بپردازند.	۱. همگرایی محلی ضعیف ۲. پیدا کردن جواب بهینه بدون اتمام تمام نقاط جستجو	۱. همگرایی سریع در رسیدن به جواب بهینه ۲. دقت در برازندگی جواب‌های بهینه	در این الگوریتم از بین همه راه حل‌ها، راه حلی که از بین دو معیار هزینه و زمان، بهینه‌ترین باشد در خروجی حاصل می‌گردد.	بهینه-سازی اجتماع ذرات
در هر تکرار از الگوریتم، زنبورهای مصنوعی با انجام جستجوهای تصادفی در اطراف جواب‌های بدست آمده در تکرار قبل به یافتن جوابهای جدید می‌پردازند.	۱. گیرافتادن در بهینه محلی	۱. یافتن جواب بهینه سراسری با کمترین تعداد تکرار ۲. همگرایی سریع	در الگوریتم ABC برای گریز از بهینگی-های محلی زنبورها به بخش‌های دیگری از فضای جستجو جهش پیدا می‌کنند و سپس در آنجا به جستجوی جواب بهینه می‌پردازند و این عمل را تا رسیدن به یک بهینه سراسری تکرار نمایند.	کلونی زنبور مصنوعی
در الگوریتم AIS همانند الگوریتم ژنتیک جواب‌ها به طور تصادفی توسط الگوریتم تولید شده و با حرکت در دامنه مساله راه حل بهینه پیدا می‌شود.	۱. جستجوی محلی عموما از پیچیدگی بالایی برخوردار است. ۲. گیرافتادن در بهینه محلی	۱. تنوع در اجرای وظایف ۲. دقت در برازندگی جوابها ۳. همگرایی سریع	در الگوریتم AIS اگر برازندگی جمعیت اولیه در چندین تکرار متوالی به راه حل بهینه دست پیدا نکرد، تمام راه حل‌هایی که کارایی بهینه‌ای ندارند حذف می‌شوند و جمعیت جدید جایگزین می‌شود و راه حل بهینه برای مساله پیدا می‌شود.	سیستم ایمنی مصنوعی

در این الگوریتم جستجو در نواحی سراسری و محلی انجام می‌گیرد و بهینه‌ترین راه حل در خروجی نمایش داده می‌شود.	۱. جستجوی محلی عموما از پیچیدگی بالایی برخوردار است. ۲. گیرافتادن در بهینه محلی	۱. همگرایی SFLA در مقایسه با الگوریتم‌های جستجوی پراکنده و الگوریتم ژنتیک بهتر است. ۲. همگرایی سریع	در این الگوریتم تمام نواحی جستجو برای راه حل بهینه جستجو می‌شود و پیدا کردن راه حل بستگی به نواحی جستجو دارد.	الگوریتم جهش قورباغه به هم ریخته
در این الگوریتم برای گریز از بهینگی-های محلی ترتیبی اتخاذ می‌شود که هنگام قرار گرفتن در بهینگی‌های محلی فضای جستجو تغییر داده شود و ارزیابی راه حل‌های قبلی بهینه‌ترین راه حل در خروجی نمایش داده شود.	۱. همگرایی محلی ضعیف ۲. عدم جستجو برای جواب بهینه در تمام نقاط جستجو	۱. همگرایی CSO در مقایسه با الگوریتم بهینه‌سازی ازدحام ذرات بهتر است. ۲. دقت در برازندگی جوابها	در این الگوریتم به طور مداوم شایستگی راه حل‌های مساله ارزیابی می‌شون و موقعیتی را که در آن بهترین برازنگی برای مساله وجود داشته باشد به عنوان جواب مساله ارائه می‌گردد.	بهینه سازی ازدحام گربه ها
یکی از محاسن FA در حل مسائل توابع پیچیده بهینه‌سازی پیوسته این است که قادر به تغییر حالت از یک محل بهینگی به محل بهینگی دیگر است. در الگوریتم کرم شب‌تاب، اگر بهترین راه حل بهینه پیدا نشود جستجو متوقف نمی‌شود و برای بدست آوردن راه حل بهینه جستجو در اطراف همسایگی نقاط مرحله قبل صورت می‌گیرد.	۱. گیرافتادن در بهینه محلی ۲. عدم دستیابی به بهینه-ترین جواب	۱. همگرایی سریع در رسیدن به جواب بهینه ۲. کاهش پیچیدگی ۳. همگرایی الگوریتم کرم شب‌تاب در مقایسه با الگوریتم‌های الگوریتم ژنتیک و الگوریتم جستجوی ممنوعه بهتر است.	در این الگوریتم عمل جستو برای بهینه-ترین راه حل به طور تصادفی در اطراف جوابهای بدست آمده در تکرارهای قبلی انجام می‌گیرد. این روش باعث می‌شود که الگوریتم حتما به جواب بهینه دست پیدا کند.	الگوریتم کرم شب-تاب
در این الگوریتم، ذرات تا انتهای فرایند جستجو، تعادل میان جستجوی سراسری و محلی را به شیوه‌ای مناسب حفظ می‌کنند.	۱. گیرافتادن در بهینه محلی	۱. کاهش زمان رسیدن به جواب بهینه ۲. تنوع در ارجرای وظایف برای رسیدن به بهینه‌ترین حالت	در هر تکرار از الگوریتم، ذرات با انجام جستجوهای تصادفی در اطراف جوابهای بدست آمده در تکرارهای قبلی به یافتن جوابهای بهینه می‌پردازند. برای این منظور میزان بهینگی جوابهای بدست آمده توسط ذرات محاسبه شده و بدین ترتیب جواب نزدیک به بهینه برای مساله بدست می‌آید.	جستجوی پراکنده
در این الگوریتم تمام همسایه‌های نزدیک به راه حل بهینه بررسی می-شوند و بهینه‌ترین حالت برای مساله در خروجی نمایش داده می‌شود.	۱. گیرافتادن در بهینه محلی ۲. پیچیدگی زیاد محاسباتی	۱. سرعت همگرایی و یافتن جواب بهینه ۲. محاسبه دقیق حالتهای همسایه و بررسی زمان دقیق وظایف	جستجو با یک جواب اولیه شروع می‌شود و در تکرارهای مداوم همسایه‌های نزدیک به جواب ارزیابی و بهینه‌ترین حالت برای مساله انتخاب می‌شود.	جستجوی تابو

برای تهیه برنامه زمانبندی، با توجه به میزان منابع موجود و زمان عادی فعالیت‌ها، راه حل‌های زمانبندی با استفاده از الگوریتم ژنتیک مشخص می‌شوند. زمان پایان پروژه‌ها و هزینه آن برای هر کدام از راه حل‌های زمانبندی محاسبه می‌شود. در صورتی که هیچکدام از راه حل‌های زمان‌بندی ارائه شده نتوانند هر دو هدف را ارضاء کنند، با تغییر زمان فعالیت‌ها و میزان منابع موجود، راه حل‌های زمانبندی مناسب‌تری ارائه می‌شوند. در انتها از بین همه راه حل‌های زمانبندی، راه حلی که با توجه به اهمیت هرکدام از دو معیار هزینه و زمان، بهترین شرایط را داشته باشد انتخاب می‌شود. نتایج آزمایشات نشان می‌دهد که الگوریتم ترکیبی کارایی خوبی دارد.

در مرجع [۵۲] با ترکیب الگوریتم شبیه‌سازی حرارت و منطق فازی، مدلی برای زمانبندی پروژه‌ها ارائه داده‌اند. پروژه‌های نرم‌افزاری غالباً پروژه‌های بزرگ با هزینه‌های سرمایه‌گذاری بالا هستند، لذا زمانبندی دقیق آنها عامل مؤثری در اجرای موفق این پروژه‌ها از نظر زمان و هزینه است. به دلیل وجود عوامل بسیاری مانند عوامل قطعی و غیر قطعی دخیل در زمانبندی پروژه‌های نرم‌افزاری، زمانبندی آنها مسئله پیچیده‌ای بوده و حل آن با روشهای معمول زمانبندی بسیار طاقت فرسا و زمان بر خواهد بود. از طرف دیگر، در روشهای معمولی، عوامل اصلی دخیل در زمانبندی (زمان، هزینه و روابط غیر خطی بین آنها) بطور همزمان بررسی نمی‌شود و لذا برنامه زمانبندی، بهینه نخواهد بود. به منظور بررسی قابلیت مدل ترکیبی ارزیابی بر روی داده‌های j30، j60 و j120 از دیتاست PSPLIB مورد تست قرار گرفته است. در [۵۳] از ترکیب جستجوی پراکنده و الکترومغناطیس برای RCPSP با محدودیت منابع و زمان‌های انتقال منابع استفاده شده است. مدل ترکیبی بر روی داده‌های j30، j60 و j120 از دیتاست PSPLIB مورد تست قرار گرفته است. نتایج محاسباتی نشان دهنده برتری مدل ترکیبی در مقایسه با مدلهای دیگر است. در [۵۴] الگوریتم ترکیبی برای بهینه‌یابی زمانبندی پروژه‌های نرم‌افزاری ترکیب دو الگوریتم ژنتیک وجستجوی همسایگی پیشنهاد گردیده است. الگوریتم ترکیبی دارای افزایش سرعت همگرایی و کاهش قرار گرفتن در بهینه محلی است. در مقایسه با الگوریتم‌های ژنتیک و جستجوی ممنوعه، این الگوریتم کارایی قابل قبولی از خود نشان داده است. در جدول ۴ تعدادی از الگوریتم‌های فرااکتشافی و ترکیبی از لحاظ معیارهای زمانبندی پروژه‌های نرم افزاری بر روی داده‌های مسائل استاندارد موجود در پایگاه PSPLIB و مجموعه نمونه j30 مورد بررسی قرار گرفته اند.

جدول۴: مقایسه الگوریتم های فرااکتشافی و ترکیبی از لحاظ معیارهای زمانبندی پروژه
Table 4. Comparison of metaheuristic and hybrid algorithms in terms of project scheduling criteria

Maximum computation time (seconds)	Average computation time (seconds)	Mean of the standard deviation from the critical path	Mean percentage deviations from the desired limit	Algorithm
5.5	1.17	11.45	0.15	Genetic Algorithm [52]
4.17	1.15	12.75	0.35	Bee colony [55]
NR	NR	11.73	0.27	Scatter search [56]
3.2	NR	11.71	0.46	Tabu search [57]
NR	0.76	11.94	0.22	Hybrid [58]

۲-۴- یادگیری ماشین

یک روش برای حل داده‌ها با حجم زیاد استفاده از تکنیک‌های یادگیری ماشین است. تکنیک‌های یادگیری ماشین بیشتر به منظور افزایش دانش، افزایش کارآیی و تصحیح اتوماتیک خطا مورد استفاده قرار می‌گیرند. آموزش و تست داده‌ها یکی از مهمترین کاربردهای تکنیک‌های یادگیری ماشین است که امروزه بسیار مورد توجه قرار می‌گیرد.

۲-۴-۱- شبکه‌های عصبی مصنوعی

شبکه‌های عصبی مصنوعی بدون برنامه‌ریزی قبلی یاد می‌گیرند که مسائل را حل کنند. در واقع تنظیم وزنهای ورودی هر نرون باعث یادگیری کل شبکه می‌شود که این تنظیم بر اساس مدل پیاده‌سازی شده می‌تواند با ناظر یا بدون ناظر صورت پذیرد. شبکه های عصبی مصنوعی می‌توانند دارای یک یا چندین لایه مخفی باشند. مقاوم بودن، قابلیت یادگیری، قابلیت تعمیم، سرعت بالا به دلیل پردازش‌های موازی، قابلیت سازگاری با تغییرات سیستم‌ها از ویژگی‌های شبکه‌های عصبی مصنوعی هستند. در [۵۹] از شبکه‌های عصبی مصنوعی برای مساله زمانبندی پروژه‌های نرم‌افزاری استفاده شده است. ارزیابی و نتایج بر روی داده‌های j30، j60 و j120 از دیتاست PSPLIB صورت گرفته است. شبکه عصبی مصنوعی برای زمانبندی پروژه‌های نرم‌افزاری بر مبنای روال زیر کار می‌کند:

۱- لایه ورودی : در این لایه، اطلاعات اولیه که همان مقادیر مربوط به متغیرهای مستقل و وابسته زمانبندی می‌باشند، وارد شبکه می‌گردند.

۲- لایه یا لایه‌های مخفی : این لایه که بر مبنای پیچیدگی تحلیل‌ها و فاکتورهای ارزیابی وظیفه ایجاد روابط منطقی میان متغیرهای مستقل و وابسته و یافتن فرمول بین این متغیرها را بر عهده دارد.

۳- لایه خروجی: در این لایه، رابطه یا فرمول بین متغیرهای مستقل و وابسته اولیه ارائه می‌شود.

در این مدل با دریافت پروژه‌ها رابطه منطقی میان متغیرهای مستقل و وابسته کشف می‌شوند و در مرحله بعد، مقدار متغیرهای وابسته را با دقت بسیار بالا تخمین و ارائه می‌نماید. نتایج و ارزیابی نشان می‌دهد که شبکه‌های عصبی مصنوعی کارایی بهتری در مقایسه با الگوریتم‌های الگوریتم ژنتیک و الگوریتم جستجوی ممنوعه دارند.

۲-۵- محاسبات نرم

در تکنیک‌های محاسبات نرم برخلاف شیوه‌های محاسباتی کلاسیک که تمامی توان خود را به دقیق بودن و در جهت مدل نمودن کامل مسائل معطوف می‌کنند، این روش با وجود عدم دقت، حقایق جزئی و ناقص بهترین راه حل را برای مسائل می‌کنند.

۲-۵-۱- منطق فازی

برای زمانبندی پروژه‌های نرم‌افزاری باید به عدم اطمینان موجود در مورد بعضی پارامترهای پروژه مانند هزینه و زمان پروژه توجه شود. برای درنظر گرفتن عدم قطعیت مذکور از منطق فازی استفاده شده است [۶۰]. برای زمانبندی پروژه با توجه به فازی بودن زمان وظایف، محدودیت منابع و جستجو برای یافتن بهینه‌ترین برنامه زمانبندی فازی ارائه شده است. برای تهیه این نوع برنامه زمانبندی، ابتدا با توجه به میزان منابع موجود و زمان عادی وظایف، انواع برنامه‌های زمانبندی ممکن ارائه می‌گردد. زمان پایان پروژه و هزینه آن برای هر کدام از راه حل‌های زمانبندی محاسبه می‌شود. در صورتی که هیچ کدام از راه حل‌های زمانبندی پیدا شده نتوانند فاکتورهای ذکر شده را بهینه کنند، با تغییر زمان وظایف و میزان منابع موجود، برنامه‌های زمانبندی بهینه‌تری ارائه می‌گردند. در انتها از بین همه راه حل‌های زمانبندی، راه حلی که که بهینه‌ترین معیار هزینه و زمان را داشته باشد انتخاب می‌گردد. مرجع [۶۱] با استفاده از رویکرد فازی به ارائه مدلی جدید جهت محاسبه درجه بحران فعالیت و مسیر در شبکه PERT می‌پردازد. زمان انجام فعالیت‌ها در مدلهای سنتی PERT معمولا یا بصورت مقادیر قطعی بیان می‌شوند و یا به عنوان یک متغیر تصادفی از توزیع بتا و نرمال بدست می‌آیند که در هر صورت با عدم قطعیت همراه است. به منظور کاهش عدم قطعیت و افزایش آستانه و دقت زمان فعالیت‌ها از اعداد فازی استفاده شده است. بر اساس مفهوم شناوری فعالیت‌ها و تابع عضویت اعداد فازی، یک تکنیک جدید جهت محاسبه درجه بحران فعالیت و مسیرهای شبکه معرفی و بر اساس آن مسیر بحرانی فازی پروژه مشخص می‌گردد. این مدل روشی جهت مدیریت زمان تکمیل پروژه‌ها نرم‌افزاری و ترکیب تئوری مجموعه‌های فازی با PERT، است. نتایج نشان دهنده اثر بخشی مدل فازی نسبت به روش PERT است.

۳- بحث و بررسی

تصمیم‌گیری برای زمانبندی هزینه و زمان جهت اجرای فعالیت‌های پروژه در شرایط محدودیت، مساله‌ای دشوار و مهم برای اجرای پروژه‌های نرم‌افزاری کلان است. چراکه اولاً تعدد فعالیت‌ها و وابستگی آنها به یکدیگر شرایط را پیچیده می‌کند و ثانیاً وجود عدم قطعیت در تخمین زمان و هزینه اجرای فعالیت‌ها، مساله را از حالت بهینه‌سازی خارج می‌سازد. با توجه به گسترش استفاده از منابع مالی به صورت وام و یا سرمایه گذاری؛ تامین منابع مالی برای هر فعالیت، به منظور پیشبرد مناسب پروژه با در نظر گرفتن عدم قطعیت در زمان و هزینه هر فعالیت بسیار حائز اهمیت است. در جدول ۵ به بررسی و مقایسه مدل‌های کلاسیک و مدل‌های هوش مصنوعی پرداخته‌ایم.

امروزه استفاده از تکنیک‌های هوش مصنوعی برای دستیابی به راه حل بهینه در حل مسائل بهینه‌سازی رشد چشمگیری یافته است. به دلیل افزایش پیچیدگی مسائل بهینه‌سازی و ناتوانی روش‌های ریاضی برای ارائه راه حل بهینه، تکنیک‌های هوش مصنوعی راهکار مناسبی برای مسائل بهینه‌سازی هستند. روش‌های ریاضی در حل بسیاری از مسائل علمی و مهندسی مورد استفاده قرار گرفته‌اند و دامنه بسیار وسیعی از مسائل مختلف را پوشش می‌دهند اما روش‌های ریاضی با وجود کارایی دقیق، هنوز هم در حل بسیاری از مسائل بهینه‌سازی با مشکلات زیادی مواجه هستند. استفاده از الگوریتم‌های فرااکتشافی منجر به این می‌شود که در میان همه راه حل‌های ممکن، راه حلهای بهینه‌تر انتخاب شوند. طبق بررسی‌هایی که انجام دادیم الگوریتم‌های فرااکتشافی در حل مساله زمانبندی پروژه‌های نرم‌افزاری کارایی بالایی در نزدیک شدن به جواب‌های بهینه دارند.

جدول۵: مقایسه مدل‌های کلاسیک و مدل‌های هوش مصنوعی
Table 5. Comparison of classic models and artificial intelligence models

نام رویکرد	ایده اصلی	مزایا	معایب	توضیحات
مدل‌های کلاسیک	در مدل‌های کلاسیک پروژه‌ها بر مبنای نمودار و روند جریان فعالیت‌ها نمایش داده می‌شوند.	۱. ارتباط بین فعالیت‌ها را با مراحل ترسیم تسهیل می‌کنند. با ارائه نمودار از فعالیت‌ها، بررسی عملیلات را برای مدیران پروژه تسهیل می‌کند.	۱. مدل‌های کلاسیک به شدت به وجود خطا در داده‌ها حساس می‌باشند. ۲. در مدل‌های کلاسیک، در صورت وجود نقص در داده‌ها دیگر امکان استفاده از داده-های ناقص برای تحلیل آماری وجود ندارد و در نتیجه حجم نمونه‌های مورد مطالعه به شدت کاهش می‌یابد. ۳. مدل‌های کلاسیک از درک و یافتن روابط ریاضی میان داده‌ها کاملا عاجز هستند. ۴. با توجه به حجم محاسبات بالا به ازای افزایش تعداد فعالیت‌ها، استفاده از مدلهای کلاسیک اجتناب ناپذیر است.	روند فعالیت‌های پروژه بر مبنای نمودار انجام می‌گیرد.
الگوریتم‌های فرااکتشافی	در این الگوریتم‌ها بر مبنای جستجوی موازی و گروهی مساله مورد ارزیابی قرار می‌گیرد و بهینه‌ترین راه حل در خورجی نمایش داده می‌شود.	۱. الگوریتم‌های فرااکتشافی بدون نیاز به روابط پیچیده ریاضی ندارند به حل مسائل بهینه‌سازی می‌پردازند. ۲. الگوریتم‌های فرااکتشافی برای حل مسائل بهینه‌سازی که پارامترهای زیاد، محدودیت تصادفی و چند تابع هدف در آنها وجود داشته باشد بسیار موفقیت آمیز عمل می‌کنند. ۳. با استفاده از الگوریتم‌های فرااکتشافی می‌توان به مقدار بهینه یا نزدیک به بهینه در زمان معقول دست پیدا کرد.	۱. زمان زیادی برای همگرایی لازم است اگر تابع برازندگی مناسب نباشد مساله همگرا نخواهد شد. ۲. الگوریتم‌های فرااکتشافی از متغیرهای تصادفی در فرایند بهینه‌سازی استفاده می-کنند. بنابراین، جواب‌های بدست آمده از الگوریتم‌های فرااکتشافی به نوعی ماهیت احتمالاتی دارند و با هر بار اجرا این الگوریتم‌ها منجر به نتیجه‌ای متفاوت از نتیجه بدست آمده در مرحله قبلی می‌شوند.	الگوریتم‌های فرااکتشافی برای پیدا کردن راه‌حل بهینه از روش‌های آزمون و تکرار استفاده می‌کنند.
الگوریتم‌های ترکیبی	در این الگوریتم‌ها این امکان وجود دارد که مساله با کمترین تعداد تکرار بهینه‌ترین راه حل را پیدا کند.	۱. در الگوریتم‌های ترکیبی از بیش از یک نقطه برای برای شروع عملیات جستجو استفاده می‌شود. بنابراین عملیات جستجو کاهش چشمگیری داشته و الگوریتم زودتر به جواب بهینه دست خواهد یافت. ۲. الگوریتم‌های ترکیبی انعطاف پذیری بالایی در جستجوی مساله دارند و از این طریق حل کارا و موثر یک مسئله را به صورت بهینه‌تر میسر می‌سازند. ۳. الگوریتم‌های ترکیبی با در نظر گرفتن چندین حالت ها دارای مزایای از جمله فضای جستجوی گسترده همراه با ازادی در تصمیم گیری بین جواب ها وتفسیر انها می باشد.	۱. از لحاظ زمان پیچیدگی بالایی دارند. ۲. کاهش سرعت همگرایی	مهمترین اهداف الگوریتم‌های ترکیبی انتخاب استراتژی‌حل بهینه برای تعریف معیارهای اندازه‌گیری هزینه و زمان

تکنیک‌های یادگیری ماشین	این تکنیک‌ها ذاتا چند متغیره بوده و با استفاده از آموزش و تست سعی در تخمین هزینه و زمان دارند.	۱. بررسی روابط میان متغیرها تحت تاثیر خطا قرار نمی‌گیرد. ۲. در این تکنیک‌ها این امکان فراهم شده تا علاوه بر استفاده از داده‌های کامل، امکان تخمین موارد نقصان نیز با دقت بسیار بالا میسر باشد. ۳. در این تکنیک‌ها تعداد زیادی از متغیرهای مستقل و وابسته مورد مطالعه قرار گرفته و رابطه میان آن ها بررسی می‌گردد.	۱. قواعد یا دستورات مشخصی برای طراحی شبکه جهت یک کاربرد اختیاری وجود ندارد. ۲. دقت نتایج بستگی زیادی به اندازه مجموعه آموزش دارد. ۳. آموزش شبکه ممکن است مشکل یا حتی غیرممکن باشد. ۴. پیش‌بینی عملکرد در این تکنیک‌ها به سادگی امکان‌پذیر نیست. ۵. تعیین پارامترها به همراه آموزش دشوار است. طراحی این تکنیک‌ها نیازمند داده-های زیاد و تکرار زیاد برای آموزش است.	توانایی یادگیری مقاوم در برابر خطا
محاسبات نرم	این تکنیک‌ها راه حل‌های نادقیق و تقریبی برای حل مسائلی که از نظر محاسباتی حل آنها دشوار بوده را ارزیابی و تست می-کنند.	۱. محاسبات نرم از خاصیت عدم دقیق بودن جهت حل مسأله و پایین آوردن هزینه استفاده می‌کند. ۲. در تکنیک‌های محاسبات نرم با یک تابع نگاشت می‌توان ورودی ها را به فضای خروجی‌ها مرتبط کرد. ۳. این تکنیک‌ها می‌توانند توابع غیر خطی با هر درجه از پیچیدگی را مدل کنند.	۱. انتخاب زیاد توابع و وابستگی زیاد آنها به مجموعه‌های فازی نتایج را تحت تاثیر قرار می‌دهد.	این تکنیک‌ها غیرخطی هستند و توانایی پرداختن به مسائل غیر خطی و پیچیده را دارند.

این الگوریتم‌ها با توجه به ساختار و نوع پیچیدگی مسائل، جوابهای مسائل بهینه‌سازی را با استفاده از روش آزمودن و جستجوی نقاط بهینه مورد بررسی قرار می‌دهند و دستیابی به جواب‌های نزدیک به بهینه را فراهم می‌نمایند.

زمانبندی پروژه‌های نرم‌افزاری با استفاده از الگوریتم‌های فرااکتشافی بهبود خوبی دارند و می‌توان با قاطعیت گفت الگوریتم‌های فرااکتشافی نسبت به مدل‌های کلاسیک دقیق‌تر عمل می‌نمایند. الگوریتم‌های فرااکتشافی با تکرارهای مداوم و آموزش داده‌ها این توانایی را دارند که فاکتورهای موثر در تخمین را بهینه‌تر نمایند و مقدار هزینه و زمانبندی پروژه‌های نرم‌افزاری را به حداقل برسانند. الگوریتم‌های فرااکتشافی به طور معمول در میان همه جواب‌های ممکن، جواب‌های بهتر را انتخاب می‌کنند. هر جواب قابل قبول می‌تواند بر اساس ارزش و یا مقدار برازندگی برای مساله مشخص شود.

از مهم‌ترین ویژگی‌های شبکه‌های عصبی مصنوعی وابسته نبودن آن ها به فرضیه‌های اولیه درباره داده های ورودی است؛ به این معنا که داده‌های ورودی می‌توانند هرگونه توزیع آماری دلخواهی داشته باشند. این ویژگی مهم شبکه های عصبی امتیاز ویژه آنها در مقابل مدل‌های کلاسیک است و به آنها این توانایی را می‌دهد که به طور یکسان از انواع مختلف داده های ورودی با هر توزیع دلخواه استفاده کنند. همچنین با داشتن قابلیت‌های بسیار دیگری مانند سرعت پردازش بالا (به دلیل پردازش‌های موازی)، داشتن توان بالقوه در حل مسائلی که شبیه‌سازی آنها از طریق منطقی یا سایر روش‌ها مشکل یا غیر ممکن است، کار آمد بودن شبکه برای یادگیری و انطباق با محیط در صورت تغییر در موقعیت محیطی، عدم از کار افتادگی شبکه در صورت آسیب دیدگی قسمتی از نورون‌ها و داشتن جواب منطقی برای داده‌هادر شرایط اطمینان، علاقمندی به استفاده از شبکه‌های عصبی مصنوعی را بیشتر کرده است.

۴- نتیجه‌گیری و کارهای آینده

در این مقاله به بررسی تکنیک‌های زمانبندی پروژه‌های نرم‌افزاری پرداخته شده است. بررسی‌های انجام شده نشان می‌دهد روش‌های هوش مصنوعی کارایی بهتری در مقایسه با مدل‌های کلاسیک دارند. دقیق بودن زمانبندی پروژه‌های نرم‌افزاری باعث می‌شود که مدیران، پروژه‌های نرم‌افزاری را بر مبنای یک چهارچوب مشخص هزینه و زمانبندی نمایند. پروژه‌های نرم‌افزاری،

زمانبندی پروژه‌های نرم‌افزاری یک مرحله اساسی و مهم در آغاز پروژه‌ها به‌شمار می‌آید. زمانبندی پروژه‌های نرم‌افزاری یکی از فعالیت‌های اصلی در تصمیمات مدیریتی، هزینه و زمان است که در پروژه‌های نرم‌افزاری از جایگاه خاصی برخوردار است. بنابراین، برای تیم‌های نرم‌افزاری هزینه و زمان یک مرحله حیاتی به منظور توسعه و به اتمام رساندن پروژه‌ها در زمان موعد است. زمانبندی دقیق، شرکت‌های توسعه نرم‌افزاری را قادر می‌سازد تا قبل از پیاده‌سازی پروژه‌ها، مقدار هزینه و زمان معین را برای مدیریت بهتر پروژه‌ها مورد بررسی قرار دهند. با توجه به بررسی‌های انجام شده در این مقاله، در آینده استفاده از مدل‌های بر مبنای روش‌های هوش مصنوعی می تواند کارایی خوبی برای تخمین زمانبندی پروژه‌های نرم‌افزاری داشته باشد.

مراجع

[1] M.C. Abeyasinghe, D.J. Greenwood, and D.E. Johansen, “An Efficient Method for Scheduling Construction Projects with Resource Constraints,” *International Journal of Project Management*, vol. 19, no. 1, pp. 29-45, 2001, doi:10.1016/S0263-7863(00)00024-7.

[2] M. Lu and H. Li, “Resource Activity Critical Path Method for Construction Planning,” *Journal of Construction Engineering and Management*, vol. 129, no. 4, pp. 412-420, 2003, doi: 10.1061/(ASCE)0733-9364(2003)129:4(412).

[3] W. Herroelen, E. Demeulemeester, and B.D. Reyck, “Resource Constrained Project Scheduling: A Survey of Recent Developments,” *Computers & Operations Research*, vol. 25, no. 4, pp. 279-302, 1998, doi: 10.1016/S0305-0548(97)00055-5.

[4] E. Demeulemeester, “Minimizing Resource Availability Costs in Time-limited Project Networks,” *Management Science*, vol. 41, pp. 1590-1598, 1995, doi: 10.1287/mnsc.41.10.1590.

[5] J. Zhang, X. Shen, and C. Yao, "Evolutionary Algorithm for Software Project Scheduling Considering Team Relationships," *in IEEE Access,* vol. 11, pp. 43690-43706, 2023, doi: 10.1109/ACCESS.2023.3270163.

[6] S. Zhang, X. Song, L. Shen, and L. Xu, "Complicated Time-Constrained Project Scheduling Problems in Water Conservancy Construction," *Processes,* vol. 11, no. 4, p. 1110, 2023, doi: 10.3390/pr11041110.

[7] D.G. Malcolm, J.H. Roseboom, C.E. Clark, and W. Fazar, “Application of a Technique for Research and Development Program Evaluation,” *Operate Res*, vol. 7, no. 5, pp. 646-69, 1959, doi: 10.1287/opre.7.5.646.

[8] W. Fix and K. Neumann, “Project Scheduling by Special GERT Networks,” *Computing*, vol. 23, pp. 299-308, 1979, doi:10.1007/BF02252134.

[9] J.P. Pantouvakis and O.G. Manoliadis, “A Practical Approach to Resource Constrained Project Scheduling,” *Operational Research an International Journal*, vol. 6, no. 3, pp. 299-309, 2006, doi: 10.1007/BF02941258.

[10] J. Holland, “Adaptation in Natural and Artificial Systems,” University of Michigan, Michigan, USA, April 29, 1992.

[11] J. Liu, Y. Liu, Y. Shi, and J. Li, “Solving Resource-Constrained Project Scheduling Problem via Genetic Algorithm,” *Journal of Computing in Civil Engineering*, vol. 34, no. 2, 2020, doi: 10.1061/(ASCE)CP.1943-5487.0000874.

[12] C.K. Chang, Hsin-yi Jiang, Y. Di, D. Zhu, and Y. Ge, “Time-line Based Model for Software Project Scheduling with Genetic Algorithms,” *Information and Software Technology*, vol. 50, pp. 1142-1154, 2008, doi: 10.1016/j.infsof.2008.03.002.

[13] J. Lin, L. Zhu and K. Gao, “A genetic programming hyper-heuristic approach for the multi-skill resource constrained project scheduling problem,” *Expert Systems with Applications*, vol. 140, p. 112915, 2020, doi: 10.1016/j.eswa.2019.112915.

[14] M. Pablo, “On Evolution, Search, Optimization, Genetic Algorithms and Martial Arts towards Memetic Algorithms Caltech Concurrent Computation Program,” *C3P Report*, 1989.

[15] L. Wang and J. Liu, “A Scale-Free Based Memetic Algorithm for Resource-Constrained Project Scheduling Problems,” *LNCS*, vol. 8206, pp. 202-209, 2013, doi: 10.1007/978-3-642-41278-3_25.

[16] H. F. Rahman, R. K. Chakrabortty, and M. J. Ryan, “Memetic algorithm for solving resource constrained project scheduling problems,” *Automation in Construction*, vol. 111, p. 103052, 2020, doi: 10.1016/j.autcon.2019.103052.

[17] M. Dorigo and L. M.Gambardella, “The Colony System: A Cooperative Learning Approach to the Traveling Salesman Problem,” *IEEE Transactions on Evolutionary Computation*, vol. 1, no.1, April 1997, doi: 10.1109/4235.585892.

[18] J. Xiao, X. T. Ao, and Y. Tang, “Solving Software Project Scheduling Problems with Ant Colony Optimization,” *Computers & Operations Research*, vol. 40, pp. 33-46, 2013, doi: 10.1016/j.cor.2012.05.007.

[19] W. -N. Chen and J. Zhang, "Ant Colony Optimization for Software Project Scheduling and Staffing with an Event-Based Scheduler," in *IEEE Transactions on Software Engineering*, vol. 39, no. 1, pp. 1-17, Jan. 2013, doi: 10.1109/TSE.2012.17.

[20] W. Deng, J. Xu, and H. Zhao, "An Improved Ant Colony Optimization Algorithm Based on Hybrid Strategies for Scheduling Problem," *in IEEE Access*, vol. 7, pp. 20281-20292, 2019, doi: 10.1109/ACCESS.2019.2897580.

[21] J. Kennedy and R. C. Eberhart, “Particle Swarm Optimization,” in *Proceedings of the IEEE International Conference on Neural Networks*, pp. 1942-1948, 1995, doi: 10.1109/ICNN.1995.488968.

[22] A. M. Fahmy, ‘Optimization Algorithms in Project Scheduling’, Optimization Algorithms - Methods and Applications. InTech, Sept. 21, 2016. doi: 10.5772/63108.

[23] Q. Jia and Y. Seo, “An Improved Particle Swarm Optimization for the Resource-Constrained Project Scheduling Problem,” *Int J Adv Manuf Technol*, vol. 67, pp. 2627-2638, 2013, doi: 10.1007/s00170-012-4679-x.

[24] R.-M. Chen, Ch.-L. Wub, Ch.-M. Wang, and Sh.-T. Lo, “Using Novel Particle Swarm Optimization Scheme to Solve Resource-Constrained Scheduling Problem in PSPLIB,” *Expert Systems with Applications*, vol. 37, no. 3, pp. 1899-1910, 2010, doi: 10.1016/j.eswa.2009.07.024.

[25] H. Zhang, X. Li, Heng Li, and F. Huang, “Particle Swarm Optimization-based Schemes for Resource-Constrained Project Scheduling,” *Automation in Construction*, vol. 14, pp. 393-404, 2005, doi: 10.1016/j.autcon.2004.08.006.

[26] G. Koulinas, L. Kotsikas, and K. Anagnostopoulos, “A particle swarm optimization based hyper-heuristic algorithm for the classic resource constrained project scheduling problem,” *Information Sciences*, vol. 277, 2014, pp. 680-693, doi: 10.1016/j.ins.2014.02.155.

[27] Y. Zhang, X. Hu, X. Cao, and Ch. Wu, “An efficient hybrid integer and categorical particle swarm optimization algorithm for the multi-mode multi-project inverse scheduling problem in turbine assembly workshop,” *Computers & Industrial Engineering*, vol. 169, 2022, doi: 10.1016/j.cie.2022.108148.

[28] D. Karaboga, “An Idea Based on Honeybee Swarm for Numerical Optimization,” *Technical Report TR06, Erciyes University, Engineering Faculty, Computer Engineering Department*, 2005.

[29] B. Crawford, R. Soto, F. Johnson, M. Vargas, S. Misra, and F. Paredes, "A Scheduling Problem for Software Project Solved with ABC Metaheuristic," *Computational Science and Its Applications. ICCSA* 2015. Lecture Notes in Computer Science, vol. 9158, Springer, Cham, doi: 10.1007/978-3-319-21410-8_48.

[30] N. Nouri, S. Krichen, T. Ladhari, and P. Fatimah, "A discrete artificial bee colony algorithm for resource-constrained project scheduling problem," *5th International Conference on Modeling, Simulation and Applied Optimization (ICMSAO)*, Hammamet, Tunisia, 2013, pp. 1-6, doi: 10.1109/ICMSAO.2013.6552557.

[31] Y.J. Shi, F.Zh. Qu, W. Chen, and B. Li, "An Artificial Bee Colony with Random Key for Resource-Constrained Project Scheduling," *LNCS*, vol. 6329, pp. 148-157, Springer, 2010, doi: 10.1007/978-3-642-15597-0_17.

[32] H. Li, X. Li, and L. Gao, "A discrete artificial bee colony algorithm for the distributed heterogeneous no-wait flowshop scheduling problem," *Applied Soft Computing*, vol. 100, p. 106946, 2021, doi: 10.1016/j.asoc.2020.106946.

[33] J. Farmer, N. Packard, and A. Perelson, "The Immune System, Adaptation and Machine Learning," *Physica D*, vol. 2, pp. 187-204, 1986, doi: 10.1016/0167-2789(86)90240-X.

[34] M. Mobini, Z. Mobini, and M. Rabbani, "An Artificial Immune Algorithm for the Project Scheduling Problem under Resource Constraints," *Applied Soft Computing*, vol. 11, pp. 1975-1982, 2011, doi: 10.1016/j.asoc.2010.06.013.

[35] R. Agarwal, M. K. Tiwari, and S. K. Mukherjee, "Artificial Immune System Based Approach for Solving Resource Constraint Project Scheduling Problem," *Int J Adv Manuf Technol* , vol. 34, pp. 584-593, 2007, doi: 10.1007/s00170-006-0631-2.

[36] M. Eusuff, K. Lansey, and F. Pasha, "Shuffled Frog-Leaping Algorithm: A Memetic Meta-heuristic for Discrete Optimization," *Engineering Optimization*, vol. 38, no. 2, pp. 129-54, 2006, doi: 10.1080/03052150500384759.

[37] C. Fang and L. Wang, "An Effective Shuffled Frog-Leaping Algorithm for Resource-Constrained Project Scheduling Problem," *Computers & Operations Research*, vol. 39, pp. 890-901, 2012, doi: 10.1016/j.cor.2011.07.010.

[38] O.A. Ameen *et al.*,"Application of shuffled frog-leaping algorithm for optimal software project scheduling and staffing," in *Innovative Systems for Intelligent Health Informatics. IRICT 2020. Lecture Notes on Data Engineering and Communications Technologies*, pp. 293-303. Cham: Springer International Publishing, 2020, doi: 10.1007/978-3-030-70713-2_28.

[39] Sh.-Ch. Chu and P.-W. Tsai "Computational Intelligence Based on the Behavior of Cats," *International Journal of Innovative Computing, Information and Control*, vol. 3, no. 1, 2007.

[40] L. Xu and W. Hu, "Cat swarm Optimization-based Schemes for Resource-Constrained Project Scheduling," *Applied Mechanics and Materials Vols*. 220-223, pp. 251-258, 2012, doi: 10.4028/www.scientific.net/AMM.220-223.251.

[41] X.S. Yang, *Nature-Inspired Meta-heuristic Algorithms*. Luniver Press, 2008.

[42] P. Sanaei, R. Akbari, V. Zeighami, and S. Shams, "Using Firefly Algorithm to Solve Resource Constrained Project Scheduling Problem," *Proceedings of Seventh International Conference on Bio-Inspired Computing: Theories and Applications*, *Advances in Intelligent Systems and Computing*, 2013, vol. 201, pp. 417-428, doi: 10.1007/978-81-322-1038-2_35.

[43] M. Dam and M. Zachariasen, "Tabu Search on the Geometric Traveling Salesman Problem," *Meta-Heuristics: Theory and Applications*, pp. 571-587, 1995, doi: 10.1007/978-1-4613-1361-8_34.

[44] O. Atli and C. Kahraman, “Fuzzy Resource-Constrained Project Scheduling Using Taboo Search Algorithm,” *International Journal of Intelligent Systems*, vol. 27, pp. 873-907, 2012, doi: 10.1002/int.21552.

[45] Y. Ma, Zh. He, N. Wang, and E. Demeulemeester “D. Erik. Tabu search for proactive project scheduling problem with flexible resources,” *Computers & Operations Research*, vol. 153, no. 2, p. 106185, 2023, doi: 10.1016/j.cor.2023.106185.

[46] F. Glover, “A Template for Scatter Search and Path Relinking,” *Lecture Notes in Computer Science*, vol. 1363, pp. 13-54. 1997, doi: 10.1007/BFb0026589.

[47] T. Sari, V. Cakir, S. Kilic, and E. Ece, "Evaluation of scatter search and genetic algorithm at resource constrained project scheduling problems," *15th IEEE International Conference on Intelligent Engineering Systems*, Poprad, Slovakia, 2011, pp. 127-130, doi: 10.1109/INES.2011.5954732.

[48] M. D. Mahdi Mobini, M. Rabbani, M. S. Amalnik, J. Razmi, and A. R. Rahimi-Vahed, “Using an enhanced scatter search algorithm for a resource-constrained project scheduling problem,” *Soft Computing*, vol. 13, pp. 597-610, 2009, doi: 10.1007/s00500-008-0337-5.

[49] D. S. Yamashita, V.A. Armentano, and M. Laguna, “Scatter Search for Project Scheduling with Resource Availability Cost,” *European Journal of Operational Research*, vol. 169, pp. 623-637, 2006, doi: 10.1016/S1568-4946(02)00065-0.

[50] K.W. Kima, M. Genb, and G. Yamazaki, “Hybrid Genetic Algorithm with Fuzzy Logic for Resource-Constrained Project Scheduling,” *Applied Soft Computing*, vol. 2/3F, pp. 174-188, 2003, doi: 10.1016/j.ins.2009.11.044.

[51] W. Chen, Y.-J. Shi, H.-F Teng, X.P. Lan, and L.-C. Hu, “An Efficient Hybrid Algorithm for Resource-Constrained Project Scheduling,” *Information Sciences*, vol. 180, pp. 1031-1039, 2010, doi: 10.1016/j.ins.2009.11.044.

[52] D. Debels, B.De Reyck, R. Leus, and M. Vanhoucke, “A Hybrid Scatter Search/Electromagnetism Meta-Heuristic for Project Scheduling,” *European Journal of Operational Research*, vol. 169, no. 2, pp. 638-653, 2006, doi: 10.1016/j.ejor.2004.08.020.

[53] S. K. Shukla, Y. J. Son and M.K. Tiwari, “Fuzzy-based Adaptive Sample-Sort Simulated Annealing for Resource-Constrained Project Scheduling,” *International Journal Advanced Manuf Technol*, vol. 36, pp. 982-995, 2008, doi: 10.1007/s00170-006-0907-6.

[54] S. Proon and M. Jin, “A Genetic Algorithm with Neighborhood Search for the Resource-Constrained Project Scheduling Problem,” *Naval Research Logistics*, vol. 58, pp. 74-82, 2011, doi: 10.1002/nav.20439.

[55] V. Valls, F. Ballestín, and S. Quintanilla, “Justification and RCPSP: A technique that pays”, *European Journal of Operational Research,* vol. 165, no. 2, pp. 375–386, 2005, doi: 10.1016/j.ejor.2004.04.008.

[56] Y. j. Shi, F.Z. Qu, W. Chen, and B. Li, “An Artificial Bee Colony with Random Key for Resource-Constrained Project Scheduling, *” Life System Modeling and Intelligent Computing. ICSEE LSMS 2010. Lecture Notes in Computer Science*, vol 6329, 2010, pp. 148-157, doi: 10.1007/978-3-642-15597-0_17 .

[57] K. Nonobe and T. Ibaraki, “Formulation and tabu search algorithm for the resource constrained project scheduling problem,” *Hansen, P. (ed.) Essays and Surveys in Metaheuristics*, pp. 557–588, Kluwer Academic Publishers, Dordrecht, 2001, doi: 10.1007/978-1-4615-1507-4_25.

[58] T. seng and L.Y. Chen, “A hybrid metaheuristic for the resource-constrained project scheduling problem,” *European Journal of Operational Research,* vol. 175, no. 2, pp. 707–721, 2006, doi: 10.1016/j.ejor.2005.06.014.

[59] A. Agarwal, V.S. Jacob, and H. Pirkul, “Augmented Neural Networks for Task Scheduling,” *European Journal of Operational Research*, vol. 151, no. 3, pp. 481-502, 2003, doi: 10.1016/S0377-2217(02)00605-7.

[60] L.D. Long and A. Ohsato, "Fuzzy critical chain method for project scheduling under resource constraints and uncertainty," *International Journal of Project Management*, vol. 26, no. 6, pp. 688-698, 2008, doi: 10.1016/j.ijproman.2007.09.012.

[61] W. Huang, L. Ding, B. Wen, and B. Cao, "Project Scheduling Problem for Software Development with Random Fuzzy Activity Duration Times," *LNCS*, vol. 5552, pp. 60-69, 2009, doi: 10.1007/978-3-642-01510-6_8.

JOURNAL OF SOUTHERN COMMUNICATION ENGINEERING
ISLAMIC AZAD UNIVERSITY BUSHEHR BRANCH

E-ISSN: 2980-9231
https://sanad.iau.ir/journal/jce

https://doi.org/10.30495/jce.2023.1988692.1209

Vol. 14/ No. 56/2025

Research Article

Novel Techniques to Reduce Transformer Inrush Current Using EMTP Software

Amir Ghaedi[1*] | Reza Sedaghati[2] | Mehrdad Mahmoudian[3]

[1]Department of Electrical Engineering, Dariun Branch, Islamic Azad University, Dariun, Iran.
amir.ghaedi@iau.ac.ir

[2]Department of Electrical Engineering, Beyza Branch, Islamic Azad University, Beyza, Iran.
Reza.sedaghati@iau.ac.ir

[3]Department of Electrical and Electronic Engineering, Shiraz University of Technology, Shiraz, Iran.
m.mahmoudian@sutech.ac.ir

Correspondence
Amir Ghaedi, Associate Professor of Electrical Engineering, Dariun Branch, Islamic Azad University, Dariun, Iran.
Amir.ghaedi@miau.ac.ir

Main Subjects:
Power Transformer

Paper History:
Received: 13 June 2023
Revised: 15 July 2023
Accepted: 25 July 2023

Abstract

Transformer inrush current is a high-amplitude, non-sinusoidal transient that occurs during initial energization, leading to voltage dips, power quality degradation, and relay misoperations. Unlike fault currents, inrush currents can be identified through harmonic analysis. Mitigating their amplitude is critical to avoiding these disruptions. This paper investigates novel inrush current mitigation techniques, with a focus on core magnetic materials. By evaluating their magnetic properties, we simulate the amplitude and harmonic composition of inrush current using EMTP-RV software. A comparative analysis identifies the most effective material for suppression. Additionally, alternative mitigation methods such as controlled energization timing, residual flux regulation, and transformer pre-loading are assessed. Their impact on inrush current is modeled in EMTP-RV, offering insights into optimal suppression strategies. The study aims to improve transformer reliability and power system stability by minimizing the adverse effects of inrush current.

Keywords: Inrush Current, Transformer, Magnetic Characteristic, Energizing Time, Residual Flux.

Highlights

- Performing a complete study on the effective methods used for reducing transformer inrush current.
- Comparing various magnetic materials used as transformer cores regarding the magnitude of inrush current, core loss, and costs.
- Investigating the effect of introduced techniques on reducing inrush current by conducting studies in EMTP-RV software.

Citation: A. Ghaedi, R. Sedaghati, and M. Mahmoudian, "Novel Techniques to Reduce Transformer Inrush Current Using EMTP Software," *Journal of Southern Communication Engineering*, vol. 14, no. 56, pp. 118–139, 2025, doi:10.30495/jce.2023.1988692.1209 [in Persian].

1. Introduction

Transformer inrush current is a high-magnitude transient current that occurs during the initial cycles after energization, caused by the nonlinear magnetic characteristics of the transformer core. Due to its large amplitude and prolonged duration, this current can trigger protection system malfunctions, leading to unintended transformer disconnection. Beyond relay misoperation, inrush current also induces voltage dips, thermal losses, and power quality degradation. Its magnitude depends on several factors, including winding resistance, switching angle, and residual flux density at the moment of energization.

Extensive research has been conducted to analyze the origins, characteristics, and mitigation of inrush current, as well as methods to distinguish it from fault currents. Key studies include:

- Evaluates the impact of inrush current on transformer protection systems. Using COMTRADE-based digital simulations, the study examines its effects on overcurrent and differential protection schemes [1].
- Investigates inrush current transients and proposes a sequential phase energization method with a neutral resistor to limit current magnitude. An analytical approach for resistor sizing is developed using nonlinear transient circuit analysis [2].
- Introduces a box dimension-based technique to differentiate inrush currents from internal faults, leveraging waveform disparities to prevent differential protection misoperation [3].
- Presents simulation and experimental results for sequential phase energization, demonstrating how a neutral grounding resistor significantly reduces inrush current [4].
- Analyzes the relationship between breaker closing voltage and inrush current magnitude, validating the method via steady-state theory and experimental data [5].
- Develops a dynamic transformer model to predict inrush current after switch-off, focusing on residual flux effects and transient overshoot [6].
- Proposes a ferromagnetic core model incorporating hysteresis, saturation, and eddy current losses. An artificial neural network estimates inrush current using experimental hysteresis data [7].
- Introduces a low-power DC-based device to reduce residual flux across multiple transformers simultaneously, verified through time-domain simulations [8].
- Proposes a normalized grille curve method to distinguish inrush currents from faults in time and frequency domains, validated with 268 tests on a YNd11 transformer [9].
- Employs error estimation to identify inrush currents by comparing actual waveforms with reference signals under varying frequency conditions [10].

This body of research advances strategies to mitigate the adverse effects of inrush currents, thereby enhancing transformer reliability and power system stability.

2. Innovations and Contributions

The key contributions of this paper are:

- **Comparative analysis of transformer core materials**, evaluating their impact on inrush current magnitude, core losses, and cost-effectiveness.
- **Comprehensive review of inrush current mitigation techniques**, providing insights into effective strategies for minimizing transient effects.

3. Materials and Methods

This study investigates the influence of different magnetic core materials on transformer inrush current. Magnetic materials are critical in the design of electromagnetic devices (e.g., transformers, generators, and motors), as they directly affect performance, efficiency, and cost.

To optimize design parameters such as cost, size, and operational quality, we analyze the following core materials:

- **Silicon steel** (common, cost-effective)
- **Nickel-iron alloys** (high permeability, low losses)
- **Cobalt iron** (high saturation flux density)
- **Amorphous metallic alloys** (low core losses)
- **Ferrites** (high resistivity, suitable for high frequencies)
- **Moly-perm alloy** (stable permeability)
- **Sendust** (low magnetostriction)
- **Iron powder cores** (distributed air gaps, reduced eddy currents)

The study employs simulations and theoretical modeling to assess each material's impact on inrush current characteristics. The model of the transformer in the EMTP-RV software is represented in Figure I.

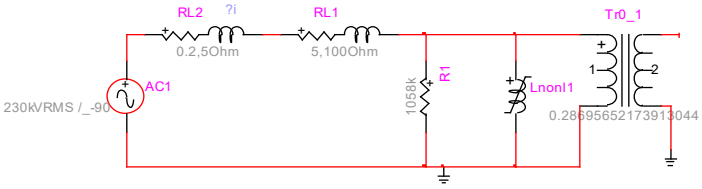


Figure I. The model of the transformer in the EMTP-RV software

The algorithm for Mitigating Transformer Inrush Current is listed as follows:

1. Problem Definition:
   - Identify challenges posed by transformer inrush current (e.g., mechanical stress, relay mis-operation).
2. Candidate Material Selection:
   - List common core materials (e.g., silicon steel, nanocrystalline, amorphous, ferrite).
3. Key Parameter Comparison:
   - Inrush Current: Simulate or experimentally measure inrush current for each material.
   - Core Losses: Quantify hysteresis and eddy current losses.
   - Cost: Analyze production and maintenance costs.
4. Material Optimization:
   - Develop a decision matrix with weighted parameters (e.g., 50% inrush, 30% losses, 20% cost).
5. Identify Critical Parameters:
   - Residual flux, switching angle, winding resistance, and core design.
6. Simulation/Experimentation:
   - Evaluate parameter impact using tools like EMTP or ANSYS.
7. Review Existing Methods:
   - Pre-fluxing: Demagnetize the core before energization.
   - Controlled Switching: Synchronize voltage phase angle.
   - Pre-insertion Resistors: Limit current transient.
   - Optimized Core Design: Air gaps or split-core techniques.
8. Method Evaluation:
   - Compare cost, technical feasibility, and effectiveness.
9. Optimal Combination:
   - Select best-performing core material + mitigation technique.
10. Validation:
    - Prototype testing or final simulation.

4. Results and Discussion

To compare different magnetic materials used as the magnetic core of the transformer from produced inrush current point of view, the peak current magnitude, the main harmonic current magnitude, the magnitude of the second and third harmonics, the DC component of the inrush currents and the damping time associated with different magnetic materials are determined and presented in table I.

Table I: The result of the simulation

material	Peak current (A)	Main harmonic (A)	2nd harmonic (A)	3rd harmonic (A)	DC current (A)	Damping time
Ferrite K 25˚C	4300	1400	100	320	400	0.4 S
Ferrite K 100˚C	3700	1480	50	190	400	0.4 S
Ferrite P&R 25˚C	3400	1120	100	240	400	0.6 S
Ferrite P&R 100˚C	4100	1550	80	270	400	0.5 S
Ferrite F 25˚C	3550	1180	100	260	400	0.5 S
Ferrite F 100˚C	4300	1700	70	250	400	0.5 S
Ferrite W&H 25˚C	3200	1190	80	190	400	0.6 S
Ferrite W&H 100˚C	4600	2100	75	180	400	0.4 S
Amorphous 2714AF	2100	750	90	100	400	0.9 S
Moly perm alloy	4000	1780	30	50	400	0.6 S
High Flux Powder	3200	1450	15	20	400	0.9 S
Sendust powder	3300	1350	30	50	400	0.6 S
75 perm powder	3150	1260	40	60	400	0.9 S

As can be seen from the figures and the table, among different magnetic materials, amorphous 2714AF produces the inrush current with the minimum magnitude. However, the time required to damp the produced inrush current

is high in this case. Among these materials, the K-type ferrite (both at 25°C and 100°C) and W&H-type ferrite (at 100°C) exhibit lower damping times compared to the other materials. The DC components of all inrush currents are 400A. It is concluded from this analysis that the best magnetic material that can be used as a magnetic core in the transformer, based on the magnitude of the produced inrush current, is amorphous 2714AF.

5. Conclusion

This paper investigates transformer inrush current and its adverse effects, including voltage drops, protective relay malfunctions, heat losses, and power quality issues. The study examines how core material properties, switching angle, residual flux, and load conditions influence inrush current. Using EMTP-RV simulations, various magnetic core materials such as K-type, P&R-type, F-type, and W&H-type ferrites, amorphous alloy 2714AF, and powdered alloys are analyzed. Results show that amorphous 2714AF yields the lowest inrush current magnitude. Further analysis reveals that the cosine voltage waveform minimizes inrush current, while the sine waveform produces the worst-case scenario. Residual flux polarity also plays a critical role: opposing the initial magnetic flux reduces inrush current. Additionally, transformer loading significantly impacts damping time, and magnitude higher loads result in smaller inrush currents and faster decay. Thus, energizing transformers under load is recommended for effective inrush current suppression. These findings provide practical insights for optimizing transformer performance and power system stability.

References

[1] L. -C. Wu, C. -W. Liu, S. -E. Chien, and C. -S. Chen, "The Effect of Inrush Current on Transformer Protection," *38th North American Power Symposium, Carbondale*, IL, USA, 2006, pp. 449-456, doi: 10.1109/NAPS.2006.359611.

[2] S.G. Abdulsalam and W. Xu, "Analytical study of transformer inrush current transients and its applications," *Proceedings International Conference on Power Systems Transients (IPST)*, Montreal, Canada, 2005, pp. 1-5.

[3] M. Gong, R. Zhang, H. Linyuan, W. Jingyu, N. Wu, "A Method for Identification of Transformer Inrush Current Based on Box Dimension," *Mathematical Problems in Engineering*, vol. 2017, pp. 1-6, 2017, doi: 10.1155/2017/2095896.

[4] Yu Cui, S. G. Abdulsalam, Shiuming Chen and Wilsun Xu, "A sequential phase energization technique for transformer inrush current reduction - Part I: Simulation and experimental results," in *IEEE Transactions on Power Delivery*, vol. 20, no. 2, pp. 943-949, April 2005, doi: 10.1109/TPWRD.2004.843467.

[5] W. Xu, S. G. Abdulsalam, Yu Cui, and Xian Liu, "A sequential phase energization technique for transformer inrush current reduction - Part II: theoretical analysis and design guide," in *IEEE Transactions on Power Delivery*, vol. 20, no. 2, pp. 950-957, April 2005, doi: 10.1109/TPWRD.2004.843465.

[6] E. Cardelli, A. Faba and F. Tissi, "Prediction and Control of Transformer Inrush Currents," in *IEEE Transactions on Magnetics*, vol. 51, no. 3, pp. 1-4, March 2015, Art no. 8400304, doi: 10.1109/TMAG.2014.2342795.

[7] J. Faiz and S. Saffari, "Inrush Current Modeling in a Single-Phase Transformer," in *IEEE Transactions on Magnetics*, vol. 46, no. 2, pp. 578-581, Feb. 2010, doi: 10.1109/TMAG.2009.2032929.

[8] B. Kovan, F. de Leon, D. Czarkowski, Z. Zabar, and L. Birenbaum, "Mitigation of Inrush Currents in Network Transformers by Reducing the Residual Flux With an Ultra-Low-Frequency Power Source," in *IEEE Transactions on Power Delivery*, vol. 26, no. 3, pp. 1563-1570, July 2011, doi: 10.1109/TPWRD.2010.2102778.

[9] JJ. Ma, Z. Wang, Q. Yang, and Y. Liu, "Identifying Transformer Inrush Current Based on Normalized Grille Curve," in *IEEE Transactions on Power Delivery*, vol. 26, no. 2, pp. 588-595, April 2011, doi: 10.1109/TPWRD.2010.2101087.

[10] B. He, Z. Xuesong, and Q.B. Zhiqian, "A new method to identify inrush current based on error estimation," *IEEE Transactions on power delivery*, vol. 21, no. 3, pp. 1163-1168, 2006, doi: 10.1109/TPWRD.2005.861337.

Declaration of Competing Interest: Authors do not have conflict of interest. The content of the paper is approved by the authors.

Author Contributions: Amir Ghaedi: Software, Methodology, Resources; **Reza Sedaghati**: Evaluation, Editing; **Mehrdad Mahmoudian**: Supervision, Writing, Editing.

E-ISSN: 2980-9231
https://sanad.iau.ir/journal/jce

مهندسی مخابرات جنوب
دانشگاه آزاد اسلامی واحد بوشهر

https://doi.org/10.30495/jce.2023.1988692.1209

مقاله پژوهشی

تکنیک‌های نوین برای کاهش جریان هجومی ترانسفورماتور با استفاده از نرم‌افزار EMTP

امیر قائدی*[۱] | رضا صداقتی[۲] | مهرداد محمودیان[۳]

[۱]گروه برق، واحد داریون، دانشگاه آزاد اسلامی، داریون، ایران.
amir.ghaedi@iau.ac.ir

[۲]گروه برق، واحد بیضا، دانشگاه آزاد اسلامی، بیضا، ایران.
Reza.sedaghati@iau.ac.ir

[۳]گروه مهندسی برق و الکترونیک، دانشگاه صنعتی شیراز، شیراز، ایران..
m.mahmoudian@sutech.ac.ir

نویسنده مسئول:
*امیر قائدی، دانشیار، گروه برق، واحد داریون، دانشگاه آزاد اسلامی، داریون، ایران.
Amir.ghaedi@miau.ac.ir

موضوع اصلی:
ترانسفورماتور قدرت

تاریخچه مقاله:
تاریخ دریافت: ۲۳ خرداد ۱۴۰۲
تاریخ بازنگری: ۲۴ تیر ۱۴۰۲
تاریخ پذیرش: ۳ مرداد ۱۴۰۲

چکیده:

جریان هجومی در ترانسفورماتور، جریانی با دامنه بالا و غیر سینوسی است که در سیکل‌های اولیه پس از برق‌دار شدن ترانسفورماتور رخ می‌دهد. این جریان ممکن است باعث بروز مشکلاتی در سیستم قدرت مانند افت ولتاژ، تلفات حرارتی، کاهش کیفیت توان و عملکرد نادرست رله‌های حفاظتی شود. برای جلوگیری از عملکرد نادرست سیستم حفاظتی، محتوای هارمونیکی این جریان برای تشخیص آن از جریان‌های خطا تحلیل می‌گردد. با این حال، تکنیک‌هایی که دامنه جریان هجومی ترانسفورماتور را کاهش می‌دهند، می‌توانند از اثرات مخرب این جریان جلوگیری کنند. به این منظور، در این مقاله مطالعه‌ای جامع بر روی تکنیک‌های نوینی که می‌توانند برای کاهش جریان هجومی ترانسفورماتور استفاده شوند، انجام شده است. ضمن تحقق این هدف، مواد مغناطیسی مختلف مورد استفاده در ساخت هسته ترانسفورماتور در نظر گرفته شده و بر اساس مشخصه مغناطیسی مربوطه، دامنه و محتوای هارمونیکی جریان هجومی تولید شده با استفاده از نرم‌افزار EMTP-RV ارزیابی می‌شود. به منظور انتخاب ماده مغناطیسی مناسب برای هسته ترانسفورماتور، مقایسه‌ای بین این مواد مغناطیسی انجام می‌شود. علاوه بر این، سایر تکنیک‌های مورد استفاده برای کاهش جریان هجومی ترانسفورماتور از جمله زمان برق‌دار شدن ترانسفورماتور، شار باقیمانده و بارگذاری ترانسفورماتور بررسی شده و تأثیر آنها بر جریان هجومی ترانسفورماتور با استفاده از نرم‌افزار EMTP-RV مدل می‌گردد.

کلید واژه‌ها: جریان هجومی، ترانسفورماتور، مشخصه مغناطیسی، زمان برق‌دار شدن، شار باقیمانده.

تازه های تحقیق:

- مطالعه کامل بر روی روشهای موثر مورد استفاده برای کاهش جریان هجومی ترانسفورماتور.
- مقایسه مواد مغناطیسی مختلف مورد استفاده به عنوان هسته ترانسفورماتور از نظر مقدار جریان هجومی، تلفات هسته و هزینه.
- بررسی تاثیر تکنیک های معرفی شده بر کاهش جریان هجومی با انجام مطالعات در نرم افزار EMTP-RV.

۱-مقدمه

جریان هجومی ترانسفورماتور، جریانی با مقدار بزرگ است که در سیکل‌های اولیه پس از برق‌دار شدن ترانسفورماتور به دلیل ویژگی غیرخطی هسته مغناطیسی ترانسفورماتور تولید می‌شود. به دلیل مقدار بزرگ و مدت زمان طولانی این جریان، سیستم حفاظتی ممکن است دچار اختلال شده و ترانسفورماتور را قطع کند. علاوه بر اختلال در سیستم حفاظتی، جریان هجومی باعث مشکلاتی مانند افت ولتاژ، تلفات حرارتی و تأثیرات بر کیفیت توان می‌شود. مقدار این جریان به پارامترهای مختلفی مانند مقاومت سیم‌پیچ، زاویه کلیدزنی و چگالی شار باقیمانده در لحظه برق‌دار شدن ترانسفورماتور بستگی دارد. برای بررسی تأثیر جریان هجومی ترانسفورماتور بر سیستم قدرت، تحقیقات زیادی برای مطالعه اساس تولید جریان هجومی، ویژگی‌های این جریان، تکنیک‌های مختلف تمایز این جریان از جریان خطا و روش‌های کاهش جریان هجومی انجام شده است.

در [۱]، تأثیر جریان هجومی بر حفاظت ترانسفورماتور ارزیابی شده است. در این مقاله، با استفاده از دو مطالعه موردی که از طریق تکنیک شبیه‌سازی دیجیتال بر اساس فایل COMTRADE انجام شده، تأثیر جریان هجومی ترانسفورماتور بر حفاظت اضافه جریان و حفاظت دیفرانسیل ترانسفورماتور بررسی شده است. مقاله [۲] به مطالعه گذرای جریان هجومی ترانسفورماتور و کاربردهای آن می‌پردازد. در این مقاله، یک رویکرد طراحی بهبود یافته برای کاهش جریان هجومی ترانسفورماتور پیشنهاد شده است. بر اساس روش پیشنهادی، هر فاز ترانسفورماتور به ترتیب برق‌دار می‌شود و از یک مقاومت خنثی برای محدود کردن جریان هجومی ترانسفورماتور سه فاز استفاده می‌شود. برای تعیین مقدار مقاومت خنثی برای عملکرد بهینه، یک رویکرد تحلیلی بر اساس تحلیل گذرای مدار غیرخطی توسعه داده شده است. در [۳]، یک روش جدید بر اساس بعد جعبه‌ای برای شناسایی جریان هجومی ترانسفورماتور به منظور جلوگیری از عملکرد نادرست حفاظت دیفرانسیل ترانسفورماتور پیشنهاد شده است. برای تمایز جریان هجومی ترانسفورماتور از جریان‌های خطای داخلی، از تفاوت اساسی در شکل موج بین دو جریان استفاده شده است. بر اساس تکنیک پیشنهادی، جریان سه فاز استخراج شده و ابعاد جعبه‌ای آن برای شناسایی جریان هجومی ترانسفورماتور محاسبه می‌شود. در [۴]، نتایج شبیه‌سازی و آزمایشگاهی مرتبط با روش برق‌دار کردن فازهای متوالی برای کاهش جریان هجومی ترانسفورماتور ارائه شده است. در این مقاله، از یک مقاومت زمین متصل به نقطه خنثی ترانسفورماتور برای عمل کردن به عنوان یک مقاومت سری با برق‌دار کردن هر فاز ترانسفورماتور به ترتیب استفاده شده و جریان هجومی ترانسفورماتور به طور قابل توجهی کاهش می‌یابد. در [۵]، تأثیر ولتاژ در دو سر کلید که باید بسته شود بر مقدار جریان هجومی ترانسفورماتور بررسی شده است. در این مقاله، بر اساس تئوری مدار حالت پایدار، ولتاژ پیشنهادی تحلیل شده و با استفاده از نتایج آزمایشگاهی، اثربخشی روش تأیید شده است. در [۶]، یک مدل دینامیکی از ترانسفورماتور برای پیش‌بینی جریان هجومی ترانسفورماتور تولید شده از تحریک اولیه دلخواه ترانسفورماتور پس از خاموش شدن توسعه داده شده است. تأثیر مدل پیشنهادی بر افزایش جریان گذرا، بر اساس پیش‌بینی شار مغناطیسی باقیمانده بررسی شده است. در [۷]، یک مدل مناسب برای هسته فرومغناطیسی ترانسفورماتور معرفی شده است تا اثرات هیسترزیس[1]، اشباع، تلفات جریان گردابی و تلفات ناهنجار ترانسفورماتور در نظر گرفته شود. در این مقاله، یک شبکه عصبی مصنوعی برای مطالعه حلقه هیسترزیس DC اصلی و منحنی مغناطیس‌شدگی اولیه ترانسفورماتور با استفاده از داده‌های تجربی پیشنهاد شده است. مدل پیشنهادی هسته فرومغناطیسی برای تخمین جریان هجومی ترانسفورماتور استفاده شده است. در [۸]، یک روش جدید برای کاهش شار مغناطیسی باقیمانده در ترانسفورماتور شبکه با استفاده از دستگاهی متشکل از منبع DC ولتاژ پایین، واحد کلیدزنی الکترونیک قدرت مناسب و یک کنترل‌کننده ساده پیشنهاد شده است. در این مقاله، با استفاده از شبیه‌سازی حوزه زمان، ثابت شده است که یک دستگاه با توان کم می‌تواند شار مغناطیسی باقیمانده همه ترانسفورماتورها را به طور هم‌زمان به میزان قابل توجهی کاهش دهد. در [۹]، یک روش جدید بر اساس منحنی شبکه نرمال شده برای تمایز جریان هجومی ترانسفورماتور از جریان‌های خطای داخلی پیشنهاد شده است. در این مقاله، دو طرح برای تمایز جریان‌های هجومی ترانسفورماتور از خطاهای داخلی در حوزه‌های زمان و فرکانس توسعه داده شده است. بر اساس ۲۶۸ نتیجه اندازه‌گیری تجربی بر روی ترانسفورماتور با اتصال YNd11، اثربخشی روش پیشنهادی تأیید شده است. در [۱۰]، یک روش جدید بر اساس تخمین خطا برای شناسایی جریان هجومی ترانسفورماتور از جریان‌های خطای

[1] Hysteresis

داخلی پیشنهاد شده است. در این مقاله، بر اساس شباهت بین موج واقعی و دو موج مرجع تحت دو شرایط فرکانسی متفاوت در هر نیم سیکل، جریان هجومی ترانسفورماتور از جریان‌های خطای داخلی تمیز داده می‌شود.

در [۱۱]، ارزیابی تحلیلی جریان هجومی ترانسفورماتور انجام شده و رویکردهای جدیدی برای کاهش جریان هجومی معرفی شده است. در این تحقیق، تأثیر پاک کردن شار باقیمانده، مقاومت پیش‌درج[1]، برق‌دار کردن بارها و کلیدزنی ناهم‌زمان بر مقدار جریان هجومی ترانسفورماتور بررسی شده است. مقاله [۱۲] تأثیر جریان هجومی بر عملکرد نادرست حفاظت فیدر خازنی در پست ۶۳ کیلوولت شاهرود را بررسی می‌کند. در این تحقیق، تأثیر مؤلفه هارمونیک فرکانس بالا و همچنین مقادیر مؤلفه DC میرا شونده جریان هجومی ترانسفورماتور که باعث اشباع هسته ترانسفورماتور می‌شود بر عملکرد نادرست برخی رله‌های حفاظتی در سیستم قدرت ارزیابی شده است. مقاله [۱۳] مدل‌سازی هارمونیکی جریان هجومی تولید شده در ترانسفورماتور قدرت نوع هسته‌ای را مطالعه می‌کند. در این مقاله، از تبدیل هارتلی[2] برای مطالعه اثرات هارمونیکی جریان هجومی ترانسفورماتور در حوزه فرکانس استفاده شده است. در [۱۴]، تکنیک برق‌دار کردن نرم ترانسفورماتور بر اساس روش تخمین زمان شیب‌دار برای کاهش مقدار جریان هجومی استفاده شده است. در این مقاله، یک تکنیک تخمین زمان شیب‌دار مبتنی بر مدل می‌تواند برای شناسایی حداقل زمان شیب ولتاژ مناسب برای مجموعه‌ای از محدودیت‌ها، مدل ترانسفورماتور و پیکربندی شبکه استفاده شود. در [۱۵]، یک استراتژی جدید بر اساس فناوری پیش‌مغناطیس کردن و کلیدزنی کنترل شده برای کاهش مقدار جریان هجومی ترانسفورماتور سه فاز پیشنهاد شده است. در این مقاله، با ساخت یک مدل مدار مغناطیسی معادل ترانسفورماتور قدرت سه فاز با ظرفیت بالا با ساختار هسته جهانی، معادلات تحلیلی شار مغناطیسی در هر مرحله از روش پیشنهادی تعیین شده است. مقاله [۱۶] جریان هجومی ترانسفورماتور قدرت را از طریق تحلیل اشباع هسته مدل می‌کند. در این تحقیق، یک رویکرد جدید بر اساس تحلیل شار هسته برای توسعه یک مدار معادل برای هسته اشباع شده ترانسفورماتور از طریق تبدیل مدار ولتاژ–جریان متعارف استفاده شده است. در جدول ۱، روش‌های استفاده شده در مقالات مرور شده خلاصه شده است.

جدول ۱: روش‌های بررسی شده در مقالات مطالعه شده

Table 1. The methods used in reviewed papers

References	Applied Methods	Research Gap
[1]-[3]-[12]-[20]-[25]	Study the impact of transformer inrush current on protection devices	Mitigation techniques are not considered.
[2]-[4]-[19]-[23]	Reducing inrush current by energizing each phase in sequence and using a natural resistor	The impact of core material, time of switch closing, residual flux, and connected load on inrush current is not considered.
[5]	The impact of breaker voltage on inrush current is studied.	The impact of core material, residual flux, and connected load on inrush current is not considered.
[6-7]-[13]-[16]-[22]-[24]	The transformer is modeled to determine the inrush current.	The impact of core material, time of switch closing, residual flux, and connected load on inrush current is not considered.
[9-10]-[17-18]-[21]	The methods for distinguishing the inrush current from other faults are studied.	The impact of core material, time of switch closing, residual flux, and connected load on inrush current is not considered.
[8]	The transformer inrush current is reduced by a device composed of a low-voltage DC source, a suitable power electronic switching unit, and a simple controller.	The impact of core material, time of switch closing, residual flux, and connected load on inrush current is not considered.
[11]-[15]	The transformer inrush current is mitigated by clearing residual flux, using a pre-insertion resistor, energizing loads, and asynchronous switching.	The impact of core material on the inrush current is not considered.
[14]	Soft transformer energizing technique is used to reduce the inrush current of the transformer.	The impact of core material, time of switch closing, residual flux, and connected load on inrush current is not considered.

در [۱۷]، یک رویکرد جدید بر اساس نمونه‌برداری تطبیقی و تبدیل هیلبرت برای تمایز جریان هجومی و خطاهای داخلی در ترانسفورماتورهای قدرت پیشنهاد شده است. تکنیک پیشنهادی می‌تواند برای تشخیص جریان هجومی همدردی و انواع مختلف خطاهای رخ داده در ترانسفورماتورهای قدرت در محیط پر نویز استفاده شود. مقاله [۱۸] تشخیص و طبقه‌بندی خطاهای بین مداری، جریان هجومی، خطاهای الکتریکی داخلی و خارجی رخ داده در ترانسفورماتورهای قدرت را مطالعه می‌کند. در این مقاله، از تبدیل موجک گسسته با حداکثر همپوشانی برای تشخیص جریان هجومی ترانسفورماتور قدرت در اختلالات مختلف

[1] pre-insertion resistor
[2] Hartley

استفاده شده است. در [۱۹]، روشی بر اساس رویکرد برق‌دار کردن نقطه روی موج و تغییر فاز متوالی برای کاهش جریان هجومی رخ داده در ترانسفورماتورهای سه فاز اعمال شده است. در این مطالعه، برای اندازه‌گیری جریان هجومی ترانسفورماتور، یک پیکربندی آزمایشگاهی با سیستم جمع‌آوری داده توسعه داده شده است. مقاله [۲۰] یک روش جدید بر اساس فایل‌های COMTRADE برای تولید مسیر دیفرانسیل رله‌های دیفرانسیل پیشنهاد می‌کند. در این مقاله، از رله دیفرانسیل برای تشخیص رویدادهای گذرا مانند جریان هجومی رخ داده در ترانسفورماتورهای قدرت استفاده شده است. در [۲۱]، تمایز جریان‌های هجومی و جریان‌های خطای داخلی رخ داده در ترانسفورماتور توسط الگوریتم فیلتر کالمن[1] توسعه یافته انجام شده است. در این مقاله، هنگام وقوع جریان هجومی، جریان سیم‌پیچ سمت اولیه و مقدار مطلق سیگنال باقیمانده توسط فیلتر کالمن توسعه یافته تخمین زده می‌شود. در [۲۲]، مطالعه شبیه‌سازی جریان هجومی ترانسفورماتور و سرکوب آن انجام شده است. در این مقاله، از MATLAB/PSB برای شبیه‌سازی جریان هجومی ترانسفورماتور سه فاز تحت شرایط کلیدزنی بی‌باری و رفع خطای خارجی استفاده شده است. مقاله [۲۳] روش کلیدزنی کنترل شده برای کاهش جریان هجومی رخ داده در ترانسفورماتورهای قدرت سه فاز سه ساق با هسته انباشته را پیشنهاد می‌کند. در تکنیک پیشنهادی، هم اشباع هسته و هم شار باقیمانده در نظر گرفته شده است. در [۲۴]، از سنسورهای هوشمند برای تشخیص اشباع ترانسفورماتور جریان در طول اندازه‌گیری جریان هجومی استفاده شده است. در این مقاله، از دو ویژگی حوزه زمان برای تمایز بین خطاها و جریان‌های هجومی رخ داده در ترانسفورماتورهای قدرت استفاده شده است. مقاله [۲۵] پارامترهای مؤثر بر اضافه ولتاژ و جریان هجومی در طول راه‌اندازی مبدل منبع ولتاژ مورد استفاده در سیستم جریان مستقیم فشار قوی را مطالعه می‌کند. در این مقاله، تأثیر اندازه خازن و سلف و همچنین زمان بسته شدن کلید قدرت بر اضافه ولتاژ و جریان هجومی ارزیابی شده است.

همانطور که در جدول ۱ مشاهده می‌شود، در تحقیقات گذشته، تأثیر مواد مختلف مورد استفاده به عنوان هسته‌های ترانسفورماتور بر مقدار جریان هجومی ترانسفورماتور در نظر گرفته نشده است. علاوه بر این، مطالعه کاملی که تمام روش‌های مؤثر مورد استفاده برای کاهش جریان هجومی ترانسفورماتور را در نظر بگیرد، انجام نشده است. با توجه به مشکلات ناشی از جریان هجومی ترانسفورماتور، این مقاله رویکردهای جامعی را برای کاهش جریان هجومی ترانسفورماتور پیشنهاد می‌کند. در ابتدا، مواد مورد استفاده به عنوان هسته مغناطیسی مطالعه و بر اساس جریان هجومی تولید شده توسط آن‌ها مقایسه می‌شوند. این مقایسه می‌تواند برای انتخاب یک ماده مغناطیسی مناسب به عنوان هسته مغناطیسی برای کاهش جریان هجومی استفاده شود. سپس، سایر پارامترهای مؤثر بر جریان هجومی ترانسفورماتور مطالعه شده و تأثیر آن‌ها بر جریان هجومی ترانسفورماتور بررسی می‌شود. بنابراین، مشارکت‌های اصلی این مقاله عبارتند از:

- مقایسه مواد مغناطیسی مختلف مورد استفاده به عنوان هسته‌های ترانسفورماتور از نظر مقدار جریان هجومی، تلفات هسته و هزینه‌ها.
- انجام یک مطالعه کامل بر روی روش‌های مؤثر مورد استفاده برای کاهش جریان هجومی ترانسفورماتور.

با توجه به هدف این مقاله، سازماندهی آن به شرح زیر خواهد بود: در بخش دوم، جریان هجومی ترانسفورماتور توضیح داده شده و موارد مختلف بیان می‌شود. در بخش سوم، مواد مغناطیسی مختلف مورد استفاده به عنوان هسته‌های مغناطیسی معرفی می‌شوند. جریان‌های هجومی تولید شده مرتبط با مواد مغناطیسی مختلف با استفاده از نرم‌افزار EMTP-RV در بخش چهارم شبیه‌سازی می‌شوند. تأثیر سایر روش‌ها بر جریان هجومی ترانسفورماتور در بخش پنجم شبیه‌سازی می‌شود. مقاله در بخش ششم خلاصه می‌شود.

۲- جریان هجومی ترانسفورماتور

در این بخش، یک ترانسفورماتور تک فاز برای مطالعه جریان هجومی تولید شده پس از برق‌دار شدن ترانسفورماتور در نظر گرفته می‌شود. در زمان برق‌دار شدن ترانسفورماتور، ولتاژ سیم‌پیچ اولیه ترانسفورماتور که توسط منبع ولتاژ تغذیه می‌شود، به صورت زیر در نظر گرفته می‌شود:

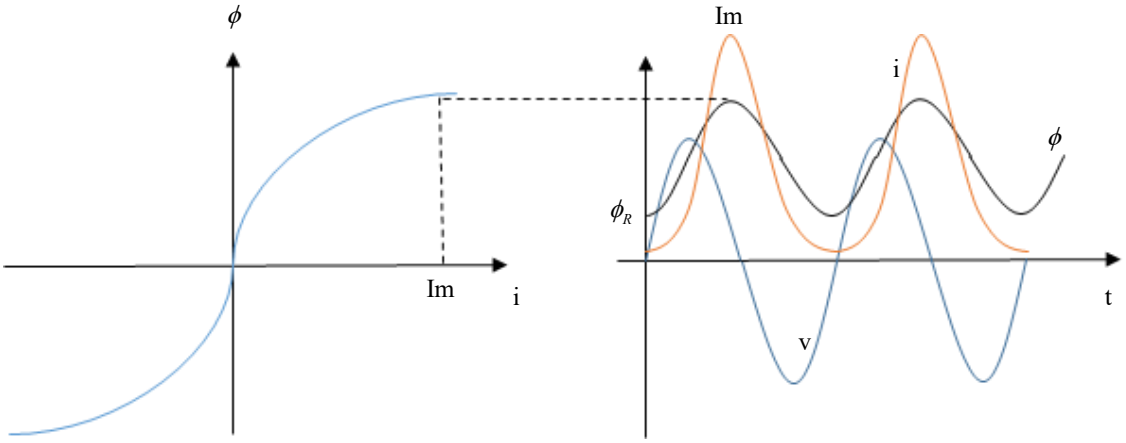
[1] Kalman filter algorithm

$$v_1(t) = V_m \sin(\omega t + \theta) \Rightarrow v_1(0) = V_m \sin\theta \quad (۱)$$

که در آن θ زاویه اولیه شکل موج ولتاژ در زمان برق‌دار شدن ترانسفورماتور است. بر اساس قانون فارادی، شار مغناطیسی هسته ترانسفورماتور را می‌توان به صورت زیر محاسبه کرد:

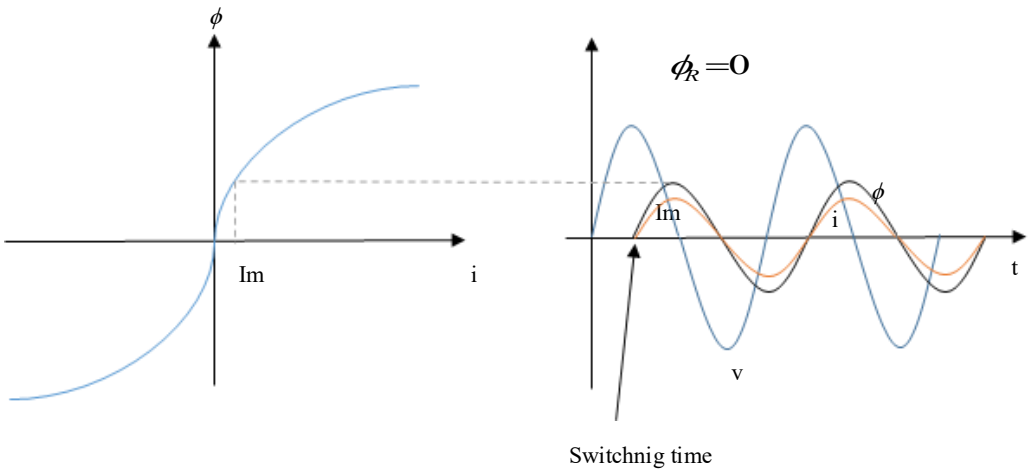
$$v_1(t) = N_1 \frac{d\phi}{dt} \Rightarrow \phi = \frac{1}{N_1}\int v_1(t)dt \quad (۲)$$

که در آن $N1$ تعداد دور سیم‌پیچ اولیه است. برای محاسبه مقدار شار مغناطیسی(ϕ) ، شار مغناطیسی اولیه یا شار مغناطیسی باقیمانده (ϕ_R) به مساحت زیر نمودار ولتاژ اضافه می‌شود، همان‌طور که در شکل ۱ نشان داده شده است.



شکل ۱: جریان هجومی با دامنه بالا در ترانسفورماتور

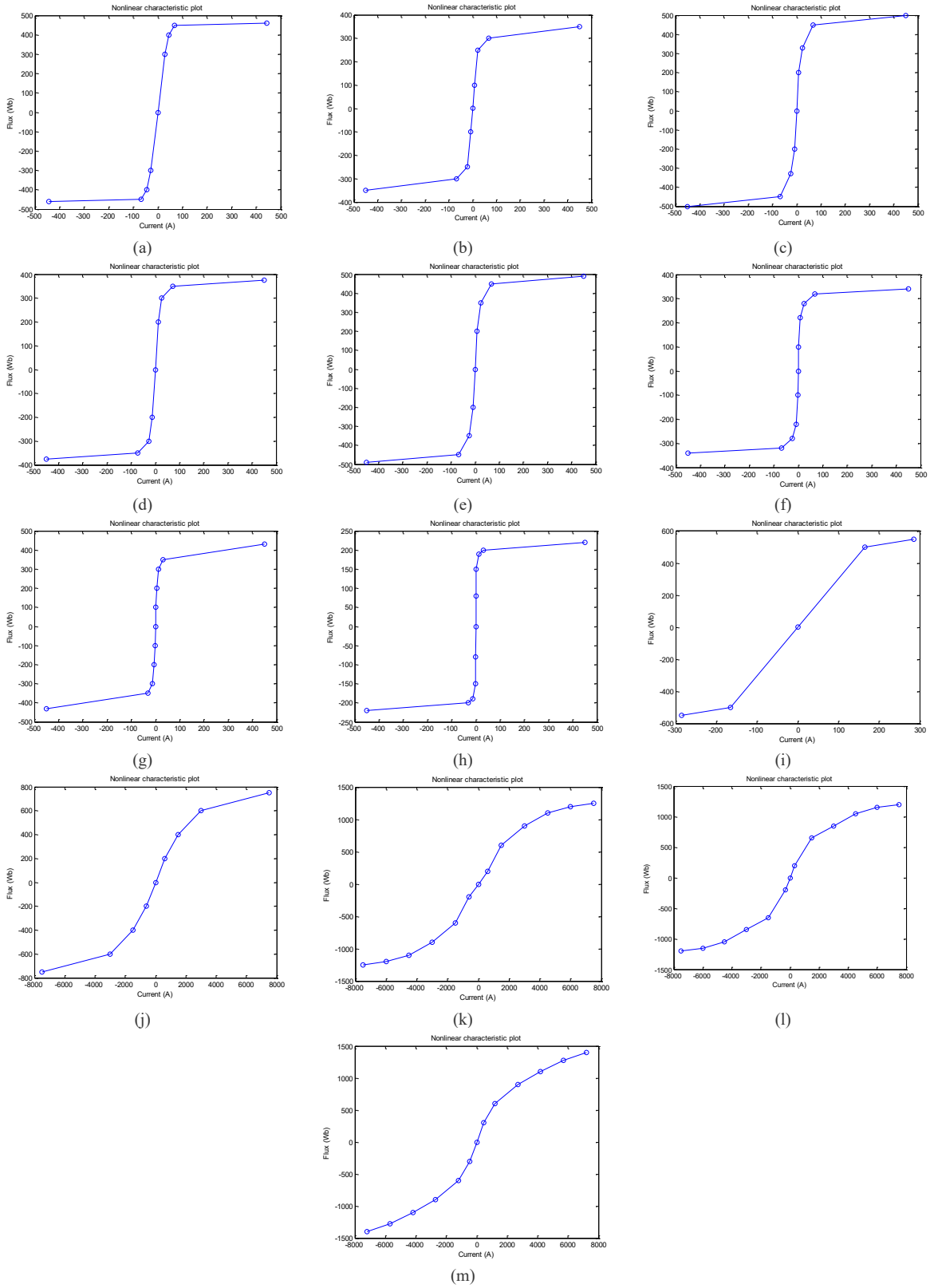
Figure 1. The high amplitude inrush current in the transformer



شکل ۲: جریان هجومی متقارن ترانسفورماتور

Figure 2. The symmetrical inrush current of the transformer

بر اساس مشخصه مغناطیسی هسته ترانسفورماتور (نمودار شار مغناطیسی-جریان تحریک) که در شکل ۱ نشان داده شده است، هر بار، بر اساس شار مغناطیسی محاسبه شده، جریان مرتبط را می‌توان تعیین کرد. همان‌طور که در شکل مشاهده می‌شود، اگر مقدار شار مغناطیسی محاسبه شده بزرگ باشد، هسته اشباع می‌شود و جریان مرتبط بسیار بزرگ خواهد بود. بنابراین، مقدار جریان هجومی بزرگ خواهد بود و به دلیل ویژگی غیرخطی منحنی مغناطیسی هسته که ناشی از اشباع هسته است، جریان هجومی غیر سینوسی بوده و حاوی جریان‌های هارمونیک است. اگر زمان کلیدزنی ترانسفورماتور مطابق شکل ۲ باشد، و شار باقیمانده ناچیز باشد، هسته در ناحیه خطی کار می‌کند و اشباع نمی‌شود و در نتیجه، مقدار جریان هجومی بزرگ نیست و حاوی جریان‌های هارمونیک نمی‌باشد.



شکل ۳: مشخصه ϕ-i مواد مغناطیسی، (a) فریت نوع K در ۲۵ درجه سانتیگراد، (b) فریت نوع K در ۱۰۰ درجه سانتیگراد، (c) فریت نوع P&R در ۲۵ درجه سانتیگراد، (d) فریت نوع P&R در ۱۰۰°C، (e) فریت نوع F در ۲۵°C، (f) فریت نوع F در ۱۰۰°C، (g) فریت نوع W&H در ۲۵°C، (h) فریت نوع W&H در ۱۰۰°C، (i) آلیاژ آمورف AF۲۷۱۴، (j) پودر آلیاژ moly perm، (k) پودر شار بالا (HF)، (l) پودر sendust و (m) پودر perm ۷۵ [26-30]

Figure 3. The ϕ-i characteristic of magnetic materials, (a) K-type ferrite at 25°C, (b) K-type ferrite at 100°C, (c) P&R-type ferrite at 25°C, (d) P&R-type ferrite at 100°C, (e) F-type ferrite at 25°C, (f) F-type ferrite at 100°C, (g) W&H-type ferrite at 25°C, (h) W&H-type ferrite at 100°C, (i) amorphous alloy 2714AF, (j) moly perm alloy powder, (k) high flux (HF) powder, (l) sendust powder and (m) 75 perm powder [26-30]

بنابراین، بر اساس رابطه ۱، اگر زاویه اولیه ولتاژ ۰ یا ۱۸۰ درجه باشد، مقدار جریان هجومی بزرگ خواهد بود؛ و اگر زاویه اولیه ولتاژ ۹۰ یا ۲۷۰ درجه باشد، جریان هجومی متقارن و مقدار آن کم است. در زمان کلیدزنی، زاویه ولتاژ ممکن است بین ۰ تا ۳۶۰ درجه باشد و بنابراین، مقدار جریان هجومی بر اساس زاویه کلیدزنی اولیه و شار باقیمانده تعیین می‌شود.

همان‌طور که از شکل‌های ۱ و ۲ مشاهده می‌شود، مقدار جریان هجومی به مشخصه مغناطیسی هسته (ϕ-i)، زاویه اولیه ولتاژ در زمان کلیدزنی و شار باقیمانده بستگی دارد. بنابراین، برای مطالعه جامع جریان هجومی ترانسفورماتور، تأثیر مشخصه مغناطیسی هسته، زمان کلیدزنی و شار باقیمانده بر جریان هجومی ترانسفورماتور بررسی می‌شود.

۳- مواد مغناطیسی مختلف

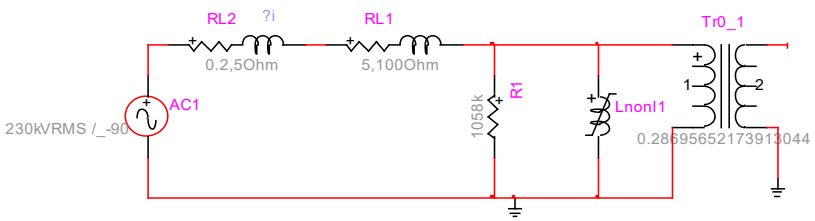
در این بخش، جریان هجومی تولید شده ترانسفورماتور مرتبط با مواد مغناطیسی مختلف مطالعه می‌شود. مواد مغناطیسی نقش کلیدی در طراحی دستگاه‌های مغناطیسی مانند ترانسفورماتورها، ژنراتورها و موتورهای الکتریکی دارند. برای دستیابی به طراحی بهینه دستگاه مغناطیسی بر اساس هزینه، اندازه و کیفیت، انواع مختلفی از مواد مغناطیسی مانند فولاد سیلیکونی، نیکل-آهن، کبالت آهن، آلیاژ فلزی آمورف[1]، فریت، آلیاژ مولی-پرم[2]، سندست[3] و پودر آهن می‌توانند استفاده شوند.

۳-۱- مشخصات مواد مغناطیسی مختلف

با توجه به کاربردهای مختلف دستگاه‌های مغناطیسی، بر اساس ویژگی‌های مطلوب از نظر مقدار چگالی شار مغناطیسی، اشباع، نفوذپذیری، رلوکتانس هسته، مقاومت و تلفات، شار باقیمانده و نقاط وادارندگی، ماده مغناطیسی مناسب انتخاب می‌شود. در این مرحله، ۱۳ ماده مغناطیسی مختلف شامل فریت نوع K در ۲۵ درجه سانتیگراد، فریت نوع K در ۱۰۰ درجه سانتیگراد، فریت نوع P&R در ۲۵ درجه سانتیگراد، فریت نوع P&R در ۱۰۰ درجه سانتیگراد، فریت نوع F در ۲۵ درجه سانتیگراد، فریت نوع F در ۱۰۰ درجه سانتیگراد، فریت نوع W&H در ۲۵ درجه سانتیگراد، فریت نوع W&H در ۱۰۰ درجه سانتیگراد، آلیاژ آمورف AF۲۷۱۴، پودر آلیاژ مولی پرم، پودر شار بالا (HF)[4]، پودر سندست و پودر ۷۵ پرم، در نظر گرفته شده و جریان‌های هجومی تولید شده این مواد ارزیابی می‌شوند. مشخصات مغناطیسی این مواد مغناطیسی در شکل ۳ نشان داده شده است.

۳-۲- نتایج شبیه‌سازی

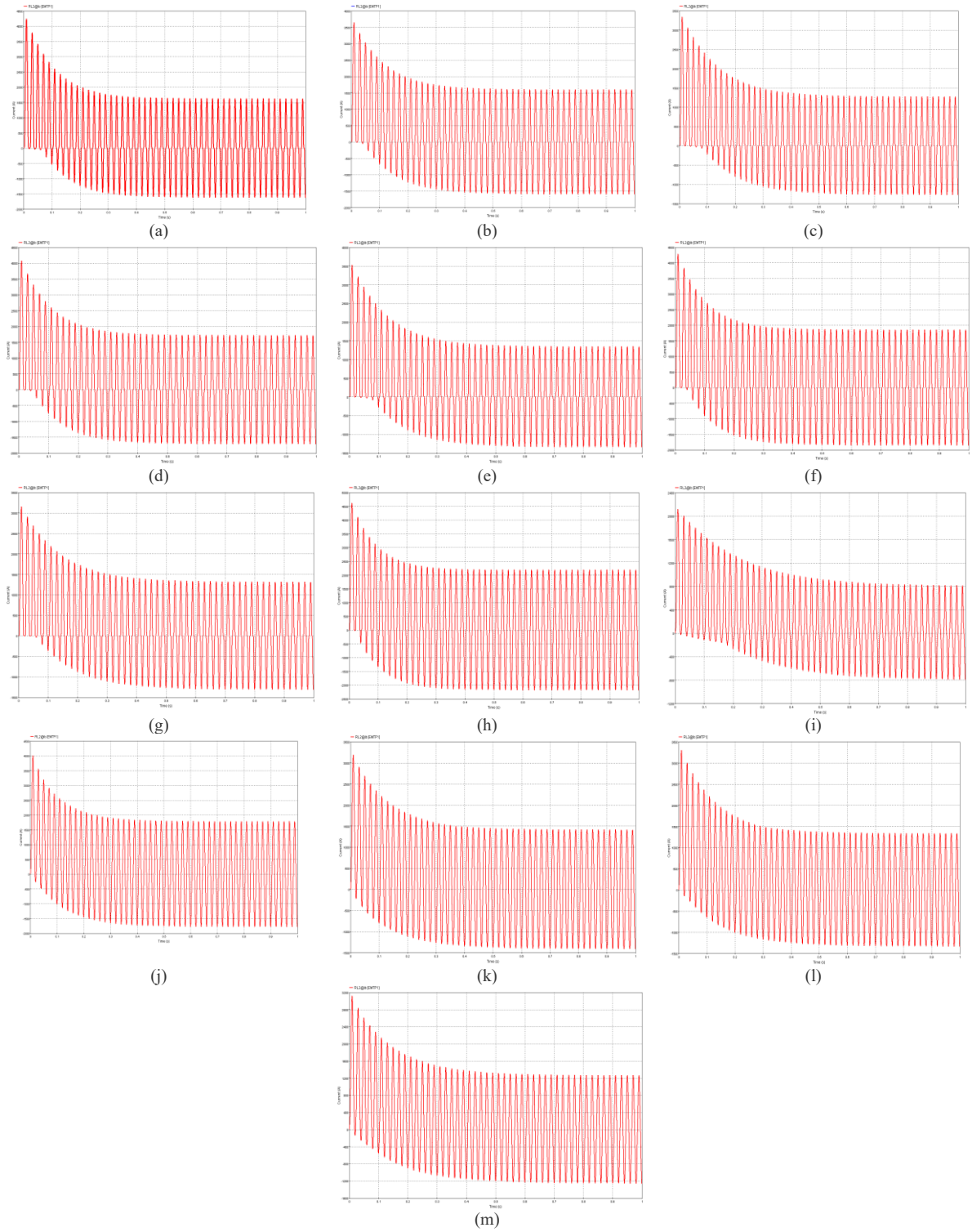
برای تعیین جریان هجومی ترانسفورماتور برای مواد مغناطیسی مختلف مورد استفاده به عنوان هسته مغناطیسی، با استفاده از نرم‌افزار EMTP-RV، یک ترانسفورماتور ۲۳۰ کیلوولت/۶۶ کیلوولت مطابق شکل ۴ مدل‌سازی شده است. منبع ولتاژ ۲۳۰ کیلوولت با فرکانس ۵۰ هرتز به ترانسفورماتور اعمال می‌شود. امپدانس از منبع ولتاژ تا ترانسفورماتور ۰/۲+j۵ اهم در نظر گرفته شده است. امپدانس سری ترانسفورماتور و مقاومت معادل هسته به ترتیب ۵+j۱۰۰ و ۱۰۵۸ اهم در نظر گرفته شده است. مشخصات غیرخطی مواد مختلف در فرآیند شبیه‌سازی در نظر گرفته شده و جریان‌های هجومی تولید شده برای شکل موج سینوسی منبع ولتاژ (بدترین حالت) و بدون شار باقیمانده تعیین و در شکل ۵ ارائه شده است. جریان‌های هارمونیک جریان‌های هجومی تولید شده با استفاده از تحلیل هارمونیک جریان‌های هجومی مربوطه تعیین و در شکل ۶ ارائه شده است.



شکل ۴: مدل ترانسفورماتور در نرم افزار EMTP-RV

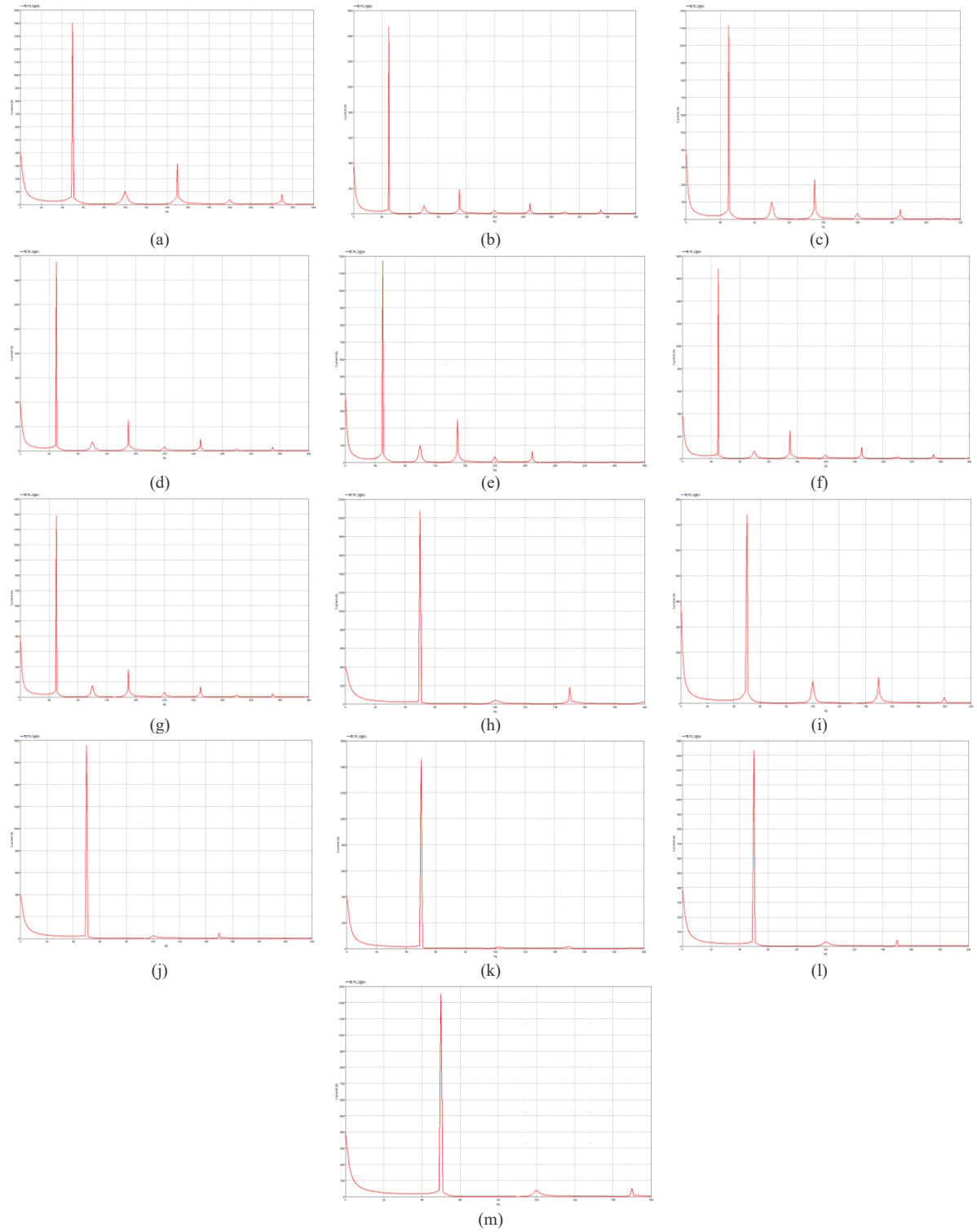
Figure 4. The model of the transformer in the EMTP-RV software

[1] amorphous
[2] moly-perm alloy
[3] sendust
[4] high flux



شکل ۵: جریان هجومی ترانسفورماتور با مواد مغناطیسی مختلف، (a) فریت نوع K در ۲۵ درجه سانتیگراد، (b) فریت نوع K در ۱۰۰ درجه سانتیگراد، (c) فریت نوع P&R در ۲۵ درجه سانتیگراد، (d) فریت نوع P&R در ۱۰۰ درجه سانتیگراد، (e) فریت نوع F در ۲۵ درجه سانتیگراد، (f) فریت نوع F در ۱۰۰ درجه سانتیگراد، (g) فریت نوع W&H در ۲۵ درجه سانتیگراد، (h) فریت نوع W&H در ۱۰۰ درجه سانتیگراد، (i) آلیاژ آمورف AF۲۷۱۴، (j) پودر آلیاژ moly perm، (k) پودر شار بالا (HF)، (l) پودر sendust و (m) پودر ۷۵ perm

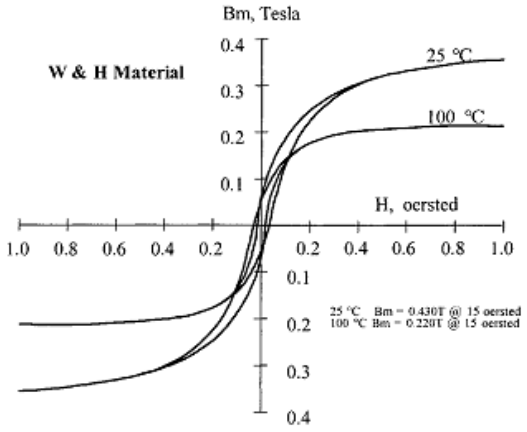
Figure 5. The transformer inrush current associated with the different magnetic materials, (a) K-type ferrite at 25°C, (b) K-type ferrite at 100°C, (c) P&R-type ferrite at 25°C, (d) P&R-type ferrite at 100°C, (e) F-type ferrite at 25°C, (f) F-type ferrite at 100°C, (g) W&H-type ferrite at 25°C, (h) W&H-type ferrite at 100°C, (i) amorphous alloy 2714AF, (j) moly perm alloy powder, (k) high flux (HF) powder, (l) sendust powder and (m) 75 perm powder



شکل ۶: هارمونیکهای جریان هجومی ترانسفورماتور با مواد مغناطیسی مختلف، (a) فریت نوع K در ۲۵ درجه سانتیگراد، (b) فریت نوع K در ۱۰۰ درجه سانتیگراد، (c) فریت نوع P&R در ۲۵ درجه سانتیگراد، (d) فریت نوع P&R در ۱۰۰ درجه سانتیگراد، (e) فریت نوع F در ۲۵ درجه سانتیگراد، (f) فریت نوع F در ۱۰۰ درجه سانتیگراد، (g) فریت نوع W&H در ۲۵ درجه سانتیگراد، (h) فریت نوع W&H در ۱۰۰ درجه سانتیگراد، (i) آلیاژ آمورف AF۲۷۱۴، (j) پودر آلیاژ moly perm، (k) پودر شار بالا (HF) ، (l) پودر sendust و (m) پودر ۷۵ perm

Figure 6. The harmonic currents of the inrush current associated with the different magnetic materials, (a) K-type ferrite at 25˚C, (b) K-type ferrite at 100˚C, (c) P&R-type ferrite at 25˚C, (d) P&R-type ferrite at 100˚C, (e) F-type ferrite at 25˚C, (f) F-type ferrite at 100˚C, (g) W&H-type ferrite at 25˚C, (h) W&H-type ferrite at 100˚C, (i) amorphous alloy 2714AF, (j) moly perm alloy powder, (k) high flux (HF) powder, (l) sendust powder and (m) 75 perm powder

همانطورکه در شکل‌ها مشاهده می‌شود، مقدار جریان هجومی تولید شده در ترانسفورماتور بالا است و پس از زمانی کمتر از ۱ ثانیه میرا می‌شود. در میان مواد مغناطیسی مختلف مورد استفاده به عنوان هسته ترانسفورماتور، آلیاژ آمورف AF۲۷۱۴، فریت نوع W&H در ۲۵ درجه سانتیگراد، پودر شار بالا (HF)، فریت نوع P&R در ۲۵ درجه سانتیگراد، پودر ۷۵ پرم و پودر سندست مقدار جریان هجومی کمتری نسبت به سایر مواد مغناطیسی دارند. علاوه بر این، فریت نوع W&H در ۱۰۰ درجه سانتیگراد، فریت نوع K در ۲۵ درجه سانتیگراد و فریت نوع F در ۱۰۰ درجه سانتیگراد مقادیر جریان هجومی بالاتری نسبت به سایر مواد مغناطیسی دارند. در این بخش، تأثیر دما بر ویژگی‌های مغناطیسی مواد استفاده شده برای هسته ترانسفورماتور ارزیابی می‌شود. شکل ۷، ویژگی B-H ماده W&H استفاده شده به عنوان هسته ترانسفورماتور را نشان می‌دهد. همانطورکه در شکل قابل مشاهده است، با افزایش دما، اشباع هسته در چگالی شار مغناطیسی کمتری رخ می‌دهد. بنابراین، با افزایش دمای هسته، مقدار جریان هجومی ترانسفورماتور افزایش می‌یابد.



شکل ۷: تاثیر دما بر مشخصه B-H مواد مغناطیسی [۲۸]

Figure 7. The impact of temperature on the B-H characteristic of magnetic materials [28]

۳-۳- بحث

برای مقایسه مواد مغناطیسی مختلف مورد استفاده به عنوان هسته مغناطیسی ترانسفورماتور از نظر جریان هجومی تولید شده، مقدار جریان پیک، مقدار جریان هارمونیک اصلی، مقدار هارمونیک‌های دوم و سوم، مؤلفه DC جریان‌های هجومی و زمان میرایی مرتبط با مواد مغناطیسی مختلف تعیین و در جدول ۲ ارائه شده است.

جدول ۲: نتایج شبیه‌سازی

Table 2. The result of the simulation

material	Peak current (A)	Main harmonic (A)	2nd harmonic (A)	3rd harmonic (A)	Dc current (A)	Damping time
Ferrite K 25°C	4300	1400	100	320	400	0.4 S
Ferrite K 100°C	3700	1480	50	190	400	0.4 S
Ferrite P&R 25°C	3400	1120	100	240	400	0.6 S
Ferrite P&R 100°C	4100	1550	80	270	400	0.5 S
Ferrite F 25°C	3550	1180	100	260	400	0.5 S
Ferrite F 100°C	4300	1700	70	250	400	0.5 S
Ferrite W&H 25°C	3200	1190	80	190	400	0.6 S
Ferrite W&H 100°C	4600	2100	75	180	400	0.4 S
Amorphous 2714AF	2100	750	90	100	400	0.9 S
Moly perm alloy	4000	1780	30	50	400	0.6 S
High Flux Powder	3200	1450	15	20	400	0.9 S
Sendust powder	3300	1350	30	50	400	0.6 S
75 perm powder	3150	1260	40	60	400	0.9 S

همان‌طور که از شکل‌ها و جدول مشاهده می‌شود، در میان مواد مغناطیسی مختلف، آمورف AF۲۷۱۴ جریان هجومی با کمترین مقدار را تولید می‌کند. با این حال، زمان لازم برای میرا شدن جریان هجومی تولید شده در این حالت بالاست. در میان این مواد، فریت نوع K (هم در ۲۵ درجه سانتیگراد و هم در ۱۰۰ درجه سانتیگراد) و فریت نوع W&H (در ۱۰۰ درجه سانتیگراد)،

زمان‌های میرایی کمتری نسبت به سایر مواد دارند. مؤلفه‌های DC تمام جریان‌های هجومی ۴۰۰ آمپر است. از این تحلیل نتیجه می‌شود که بهترین ماده مغناطیسی که می‌تواند به عنوان هسته مغناطیسی در ترانسفورماتور بر اساس مقدار جریان هجومی تولید شده استفاده شود، آمورف AF۲۷۱۴ است. برای طراحی ترانسفورماتور قدرت و انتخاب ماده مغناطیسی مناسب برای هسته ترانسفورماتور، علاوه بر مقدار جریان هجومی، تلفات هسته و هزینه هسته نیز مورد توجه قرار می‌گیرد. تلفات هسته ترانسفورماتور شامل تلفات هیسترزیس و تلفات جریان گردابی است که به چگالی شار بستگی دارد. در جدول ۳، مقدار چگالی شار (B_m) مربوط به مواد مغناطیسی مختلفی که به عنوان هسته ترانسفورماتور استفاده می‌شوند، داده شده است. تلفات هسته ترانسفورماتور در چگالی شار بالاتر بیشتر است. بنابراین، در میان مواد مغناطیسی مختلف استفاده شده به عنوان هسته ترانسفورماتور، مقدار تلفات هسته برای پودرهای با شار بالا، پودر سندوست و پودرهای ۷۵ پرم بیشتر است. با این حال، تلفات هسته ترانسفورماتور برای فریت W&H در دمای ۱۰۰۰ درجه سانتی‌گراد کمتر از سایر مواد مغناطیسی است.

جدول ۳: روش‌های بررسی شده در مقالات مطالعه شده

Table 3. The B_m and cost of magnetic materials

Material	Flux density B_m (T)
Ferrite K 25˚C	0.460
Ferrite K 100˚C	0.350
Ferrite P&R 25˚C	0.500
Ferrite P&R 100˚C	0.375
Ferrite F 25˚C	0.490
Ferrite F 100˚C	0.340
Ferrite W&H 25˚C	0.430
Ferrite W&H 100˚C	0.220
Amorphous 2714AF	0.575
Moly perm alloy	0.700
high flux Powder	1.500
Sendust powder	1.000
75 perm powder	0.950

جدول ۴ قیمت مواد مغناطیسی مختلف استفاده شده به عنوان هسته ترانسفورماتور را نشان می‌دهد. همان‌طور که در جدول دیده می‌شود، در میان مواد مغناطیسی مختلف استفاده شده برای هسته ترانسفورماتور، هزینه آلیاژ آمورف کمترین است. با این حال، در بین این مواد مغناطیسی، قیمت آلیاژ مولی پرم بیشترین است. بنابراین، برای انتخاب ماده مغناطیسی مناسب برای هسته ترانسفورماتور، مقدار جریان هجومی، تلفات هسته و قیمت باید مورد توجه قرار گیرند.

جدول ۴: روش‌های بررسی شده در مقالات مطالعه شده

Table 4. The B_m and cost of magnetic materials

Material	Price (US$/kg)
Ferrite	1.5-2.3
Amorphous	0.2-0.3
Moly perm alloy	15-45
High flux powder	2.0-2.4
Sendust powder	1-40

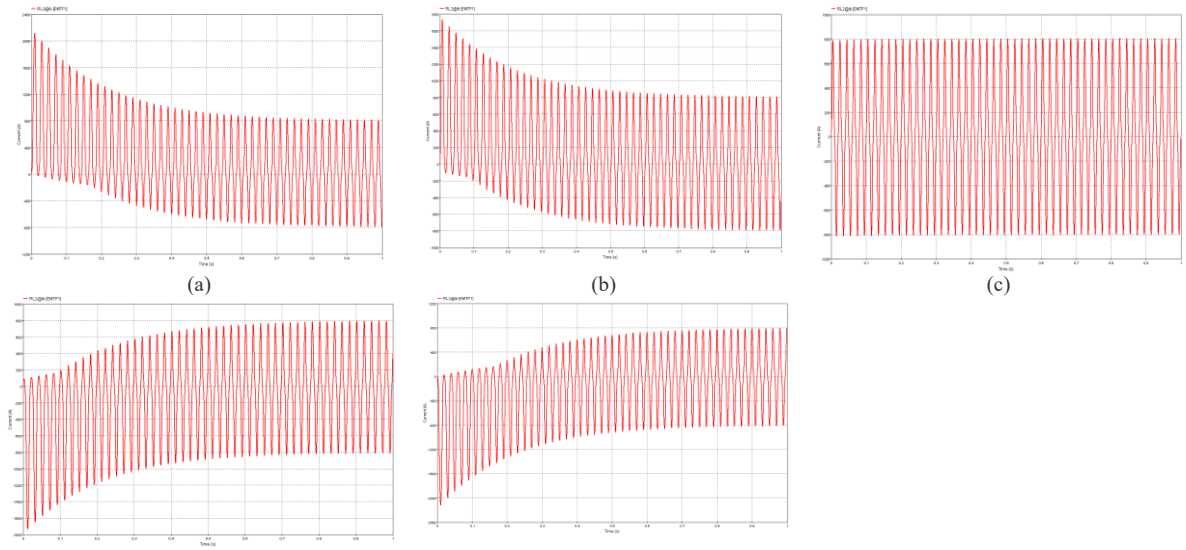
۴- سایر تکنیک‌ها برای کاهش جریان هجومی

در این بخش، سایر عوامل موثر بر کاهش جریان هجومی ترانسفورماتور، شامل زاویه ولتاژ در زمان سوئیچینگ، شار باقیمانده و بار ترانسفورماتور بررسی و با استفاده از نرم‌افزار EMTP-RV شبیه‌سازی می‌شوند.

۴-۱- زاویه ولتاژ در زمان سوئیچینگ

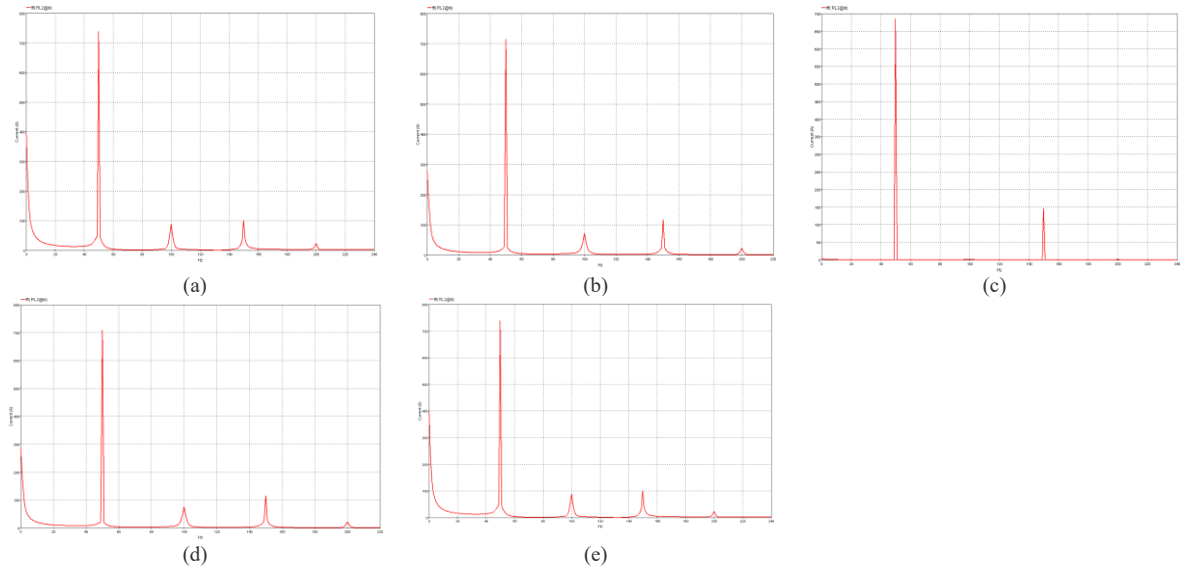
شار مغناطیسی ترانسفورماتور با افزودن شار باقیمانده به مساحت زیر نمودار ولتاژ تعیین می‌شود. اگر ولتاژ در زمان سوئیچینگ شکل موج سینوسی داشته باشد، مساحت زیر نمودار ولتاژ از زاویه صفر تا π مثبت است و بنابراین، شار تولید شده بزرگ خواهد بود. با این حال، اگر زاویه ولتاژ اولیه $\pi/2$ باشد، مساحت زیر نمودار ولتاژ از صفر تا $\pi/2$ مثبت و سپس از $\pi/2$ تا π منفی است. بنابراین، جریان هجومی مربوط به شکل موج سینوسی بسیار بزرگ (بدترین حالت) و شکل موج کسینوسی عادی و متقارن است

(بهترین حالت). در این بخش، ترانسفورماتور تحت بررسی مجهز به آمورف AF۲۷۱۴ با استفاده از نرم‌افزار EMTP-RV برای زوایای اولیه مختلف ولتاژ اعمالی به ترانسفورماتور شبیه‌سازی شده است. در این مرحله، شار باقیمانده صفر در نظر گرفته شده است. نمودارهای جریان هجومی تولید شده برای زوایای اولیه مختلف شامل ۰، ۴۵ درجه، ۹۰ درجه، ۱۳۵ درجه و ۱۸۰ درجه تعیین و در شکل ۸ ارائه شده‌اند. جریان‌های هارمونیک جریان‌های هجومی تولید شده برای زوایای اولیه مختلف ولتاژ با استفاده از تحلیل هارمونیکی جریان‌های هجومی تعیین و در شکل ۹ ارائه شده‌اند.



شکل ۸: جریان هجومی تولید شده برای زوایای اولیه (a) صفر درجه، (b) ۴۵ درجه، (c) ۹۰ درجه، (d) ۱۳۵ درجه و (e) ۱۸۰ درجه

Figure 8. The produced inrush current for initial angles of (a) 0º, (b) 45º, (c) 90º, (d) 135º and (e) 180º



شکل ۹: هارمونیکهای جریان هجومی تولید شده برای زوایای اولیه (a) صفر درجه، (b) ۴۵ درجه، (c) ۹۰ درجه، (d) ۱۳۵ درجه و (e) ۱۸۰ درجه

Figure 9. The harmonic currents associated with the inrush current for initial angles of (a) 0º, (b) 45º, (c) 90º, (d) 135º and (e) 180º

بیشینه جریان، مؤلفه DC، اندازه اصلی هارمونیک، اندازه‌های جریان هارمونیک دوم و سوم و زمان میرایی مرتبط با زوایای اولیه مختلف شکل موج ولتاژ محاسبه و در جدول ۵ ارائه شده‌اند. همانطور که از شکل‌ها و جدول مشاهده می‌شود، وقتی زوایای اولیه شکل موج ولتاژ ۰ درجه یا ۱۸۰ درجه باشند، مقدار جریان هجومی بسیار بزرگ است (۲۱۰۰ آمپر) که ۲/۶۲۵ برابر جریان هجومی مربوط به بهترین زوایای اولیه، یعنی زوایای ۹۰ درجه یا ۹۰- درجه است که مقدار جریان هجومی در این حالت‌ها ۸۰۰

آمپر است. از جدول نتیجه‌گیری می‌شود که زمان مورد نیاز برای میرایی جریان هجومی در بدترین حالت‌ها ۰/۸ ثانیه است که معادل ۴۰ سیکل با فرکانس ۵۰ هرتز است.

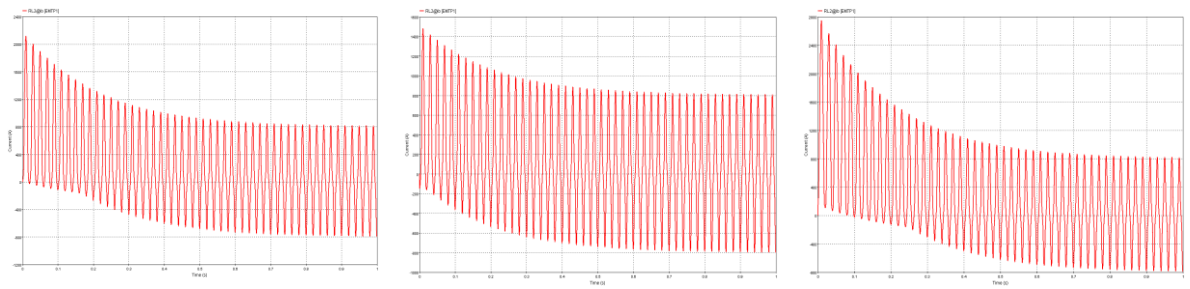
جدول ۵: نتایج شبیه سازی طبق زاویه اولیه ولتاژ

Table 5. The results of the simulation associated with the initial angle of the voltage

Switching angle (degree)	Peak current (A)	DC component (A)	Main harmonic (A)	2nd harmonic (A)	3rd harmonic (A)	Damping time (s)
0	2100	400	745	90	100	0.8
45	1750	280	710	80	110	0.6
90	800	0	680	0	150	0
135	-1750	280	710	80	110	0.6
180	-2100	400	745	90	100	0.8

۴-۲- شار باقیمانده

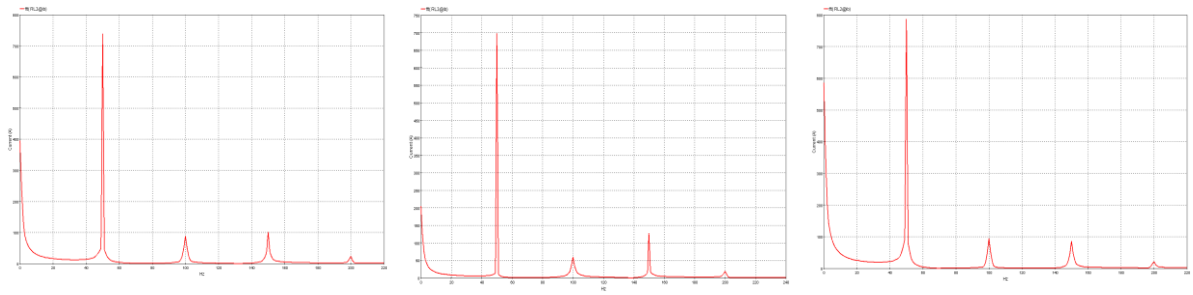
در این زیر بخش، تأثیر شار باقیمانده بر جریان هجومی ترانسفورماتور با استفاده از نرم‌افزار EMTP-RV ارزیابی می‌شود. شار مغناطیسی ترانسفورماتور با افزودن شار باقیمانده به مساحت زیر نمودار ولتاژ تعیین می‌شود. اگر علامت‌های شار باقیمانده و تغییرات شار مغناطیسی تولید شده یکسان باشند، شار مغناطیسی تولید شده به شار باقیمانده اضافه شده و شار مغناطیسی حاصل بزرگتر خواهد بود. بنابراین، جریان هجومی تولید شده بزرگتر خواهد بود. با این حال، اگر شار مغناطیسی تولید شده باعث کاهش شار باقیمانده شود، شار مغناطیسی تولید شده کمتر خواهد بود و بنابراین، جریان هجومی تولید شده کوچکتر خواهد بود. در این بخش، زاویه اولیه ولتاژ اعمال شده به ترانسفورماتور مجهز به هسته مغناطیسی آمورف AF۲۷۱۴، صفر درجه در نظر گرفته شده است. جریان‌های هجومی تولید شده برای مقادیر مختلف شار باقیمانده، شامل ۰، ۵۰۰- وبر و ۵۰۰ وبر، با استفاده از نرم‌افزار EMTP-RV تعیین شده و در شکل ۱۰ ارائه شده‌اند. جریان‌های هارمونیکی مربوط به این جریان‌های هجومی با استفاده از تحلیل هارمونیکی جریان‌های هجومی تولید شده تعیین و در شکل ۱۱ ارائه شده‌اند.



(a) (b) (c)

شکل ۱۰: جریان هجومی تولید شده برای مقادیر شار باقیمانده (a) ۰، (b) ۵۰۰- وبر، (c) ۵۰۰ وبر

Figure 10. The produced inrush current for residual flux values of (a) 0, (b) -500Wb, (c) 500Wb



(a) (b) (c)

شکل ۱۱: هارمونیک جریان هجومی تولید شده برای مقادیر شار باقیمانده (a) ۰، (b) ۵۰۰- وبر، (c) ۵۰۰ وبر

Figure 11. The harmonic currents associated with the produced inrush currents for residual flux values of (a) 0, (b) -500Wb, (c) 500Wb

بیشینه جریان، مؤلفه DC، اندازه اصلی هارمونیک، اندازه‌های جریان هارمونیک دوم و سوم و زمان میرایی مرتبط با شارهای باقیمانده مختلف محاسبه و در جدول ۶ ارائه شده‌اند. همانطور که از شکل‌ها و جدول مشاهده می‌شود، زمانی که شار باقیمانده ۵۰۰ وبر و علامت آن با علامت تغییر شار مغناطیسی یکسان است، مقدار جریان هجومی بسیار بزرگ است. با این حال، اگر شار باقیمانده ۵۰۰- وبر باشد، شار مغناطیسی تولید شده که مثبت است به شار باقیمانده اضافه می‌شود و بنابراین، شار مغناطیسی کل تولید شده کمتر از حالت‌های با شار باقیمانده صفر و ۵۰۰ وبر خواهد بود. از جدول نتیجه‌گیری می‌شود که زمان مورد نیاز برای میرایی جریان هجومی در بدترین حالت‌ها بیش از ۱ ثانیه است که بیش از ۵۰ سیکل با فرکانس ۵۰ هرتز است.

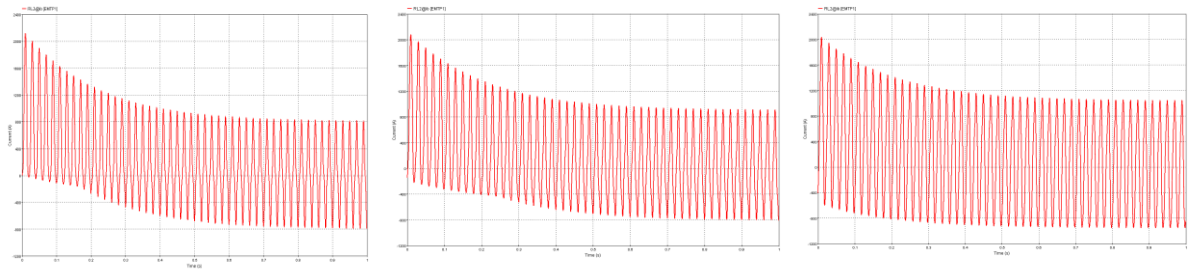
جدول۶ : نتایج شبیه سازی مطابق با شار باقیمانده

Table 6. The results of the simulation associated with the residual flux

Residual flux (Wb)	Peak current (A)	DC component (A)	Main harmonic (A)	2nd harmonic (A)	3rd harmonic (A)	Damping time (s)
0	2100	400	745	90	100	0.9
-500	1500	200	700	60	130	0.8
500	2760	590	795	95	80	>1

۴-۳- بارگذاری ترانسفورماتور

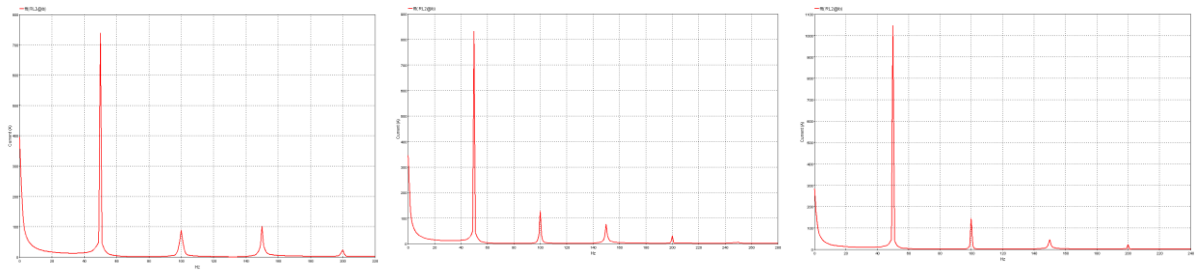
در این زیربخش، تأثیر بارگذاری ترانسفورماتور بر جریان هجومی با استفاده از نرم‌افزار EMTP-RV بررسی می‌شود. برای این منظور، سه بار مختلف شامل بدون بار، 15 MW + j7 MVAR و 45 MW + j15 MVAR بر روی ترانسفورماتور اعمال شده و جریان‌های هجومی تولید شده برای شار پسماند صفر و زاویه اولیه صفر درجه تعیین و در شکل ۱۲ ارائه شده است. با استفاده از تحلیل هارمونیک، جریان‌های هارمونیک مرتبط با این جریان‌های هجومی تعیین و در شکل ۱۳ نمایش داده شده است. جریان پیک، مؤلفه DC، دامنه هارمونیک اصلی، دامنه جریان‌های هارمونیک دوم و سوم و زمان میرایی مرتبط با بارهای مختلف ترانسفورماتور محاسبه و در جدول ۷ ارائه شده است. همانطور که از شکل‌ها و جدول مشخص است، هنگامی که ترانسفورماتور به بار متصل می‌شود، دامنه جریان هجومی تولید شده کمتر و زمان لازم برای میرایی این جریان‌ها کوتاه‌تر است.



(a) (b) (c)

شکل ۱۲: جریان هجومی تولید شده برای بارگذاریهای مختلف (a) بی باری، (b) 15MW+j7MVAR، (c) 45MW+j15MVAR

Figure 12. The produced inrush current for different transformer loading (a) no load, (b) 15MW+j7MVAR and (c) 45MW+j15MVAR



(a) (b) (c)

شکل ۱۳: هارمونیک جریان هجومی تولید شده برای بارگذاریهای مختلف (a) بی باری، (b) 15MW+j7MVAR، (c) 45MW+j15MVAR

Figure 13. The harmonic currents associated with the produced inrush currents for different transformer loading (a) no load, (b) 15MW+j7MVAR and (c) 45MW+j15MVAR

از نتایج عددی نتیجه می‌شود که هرچه بار متصل به ترانسفورماتور بزرگتر باشد، دامنه جریان هجومی تولید شده کوچکتر و زمان میرایی جریان هجومی تولید شده کوتاه‌تر خواهد بود. بنابراین، پیشنهاد می‌شود برای کاهش جریان هجومی تولید شده ترانسفورماتور، آن را تحت بار برقدار کنید.

جدول۷ : نتایج شبیه‌سازی مطابق با بارگذاری های مختلف ترانسفورماتور

Table 6. The results of the simulation associated with the different transformer loads

Load (MW,MVar)	Peak current (A)	DC component (A)	Main harmonic (A)	2nd harmonic (A)	3rd harmonic (A)	Damping time (s)
0	2100	400	745	90	100	0.9
15+7j	2050	350	840	120	80	0.8
45+15j	2020	290	1060	150	40	0.7

۵- نتیجه گیری

این مقاله به بررسی چالش‌هایی که جریان هجومی ترانسفورماتور در سیستم‌های قدرت ایجاد می‌کند، از جمله کاهش ولتاژ، اختلال در سیستم‌های محافظتی، از دست دادن حرارت و مشکلات کیفیت برق می‌پردازد. این جریان به ویژگی‌های مغناطیسی هسته، زاویه موج ولتاژ اعمال شده به ترانسفورماتور در زمان سوئیچینگ، شار باقیمانده و بار متصل به ترانسفورماتور وابسته است. به این منظور، مواد مغناطیسی مختلف شامل فریت نوع K در دماهای ۲۵ درجه سانتیگراد و ۱۰۰ درجه سانتیگراد، فریت نوع P&R در دماهای ۲۵ درجه سانتیگراد و ۱۰۰ درجه سانتیگراد، فریت نوع F در دماهای ۲۵ درجه سانتیگراد و ۱۰۰ درجه سانتیگراد، فریت نوع W&H در دماهای ۲۵ درجه سانتیگراد و ۱۰۰ درجه سانتیگراد، آلیاژ آمورف AF۲۷۱۴، آلیاژ مولی پرم، پودر با فلوکس بالا (HF)، پودر سنداست و پودر ۷۵ پرم مورد بررسی قرار می‌گیرند و بر اساس شار مغناطیسی-جریان مرتبط، جریان‌های هجومی تولید شده این مواد با استفاده از نرم‌افزار EMTP-RV محاسبه می‌شوند. از نتایج عددی استنباط می‌شود که از بین مواد مغناطیسی مختلف مورد مطالعه، از نظر مقدار جریان هجومی آلیاژ آمورف AF۲۷۱۴ بهترین ماده مغناطیسی است که می‌توان از آن برای هسته مغناطیسی استفاده کرد. سپس، تأثیر زاویه اولیه ولتاژ اعمال شده بر ترانسفورماتور بر مقدار جریان هجومی بررسی می‌شود و استنباط می‌شود که بهترین شکل موج ولتاژ که منجر به کمینه جریان هجومی می‌شود، موج کسینوس است و بدترین حالت که منجر به جریان هجومی بزرگ می‌شود، اعمال موج سینوس بر ترانسفورماتور است. عامل دیگری که در مقاله مورد بررسی قرار گرفته است، شار باقیمانده هسته است. از نتایج عددی استنباط می‌شود که برای کاهش جریان هجومی ترانسفورماتور، علامت شار باقیمانده و شار مغناطیسی تولید شده اولیه باید مخالف باشند. عامل سوم که بر جریان هجومی ترانسفورماتور تأثیر می‌گذارد، بارگذاری ترانسفورماتور است. از نتایج عددی استنباط می‌شود که هرچه بار متصل به ترانسفورماتور بزرگتر باشد، مقدار جریان هجومی تولید شده کمتر و زمان میرایی آن کوتاه‌تر خواهد بود. بنابراین، توصیه می‌شود ترانسفورماتور را تحت شرایط بارگذاری قرار دهید تا جریان هجومی تولید شده توسط آن به طور موثر کاهش یابد.

مراجع

[1] L. -C. Wu, C. -W. Liu, S. -E. Chien, and C. -S. Chen, "The Effect of Inrush Current on Transformer Protection," *38th North American Power Symposium, Carbondale*, IL, USA, 2006, pp. 449-456, doi: 10.1109/NAPS.2006.359611.

[2] S. G. Abdulsalam and W. Xu, "Analytical study of transformer inrush current transients and its applications," *The International Conference on Power Systems Transients (IPST'05)*, Montreal, Canada, 2005, pp. 1-5.

[3] M. Gong, R. Zhang, H. Linyuan, W. Jingyu, and N. Wu, "A Method for Identification of Transformer Inrush Current Based on Box Dimension," *Mathematical Problems in Engineering*, pp. 1-6, 2017, doi: 10.1155/2017/2095896.

[4] Y. Cui, S. Abdlsalam, S. Chen and W. Xu, “A sequential phase energization technique for transformer inrush current reduction-Part I: Simulation and experimental results,” *IEEE Transactions on power delivery*, vol. 21, no. 2, pp. 943-949, 2005, doi: 10.1109/PES.2004.1372857.

[5] W. Xu, S. G. Abdulsalam, Yu Cui, and Xian Liu, "A sequential phase energization technique for transformer inrush current reduction - Part II: theoretical analysis and design guide," in *IEEE Transactions on Power Delivery*, vol. 20, no. 2, pp. 950-957, April 2005, doi: 10.1109/TPWRD.2004.843465.

[6] E. Cardelli, F. Antonio, and T. Francesco, “Prediction and control of transformer inrush currents,” *IEEE Transactions on Magnetics*, vol. 51, no. 3, pp. 1-4, 2015, doi: 10.1109/TMAG.2014.2342795.

[7] J. Faiz and S. Saffar, “Inrush current modeling in a single-phase transformer,” *IEEE Transactions on Magnetics*, vol. 46, no. 2, pp. 578-581, 2010, doi: 10.1109/TMAG.2009.2032929.

[8] B. Kovan, F. Leon, D. Czarkowski, Z. Zabar, and L. Birenbaum, “ Mitigation of inrush currents in network transformers by reducing the residual flux with an ultra-low-frequency power source,” *IEEE Transactions on Power Delivery*, vol. 26, no. 3, pp. 1563-1570, 2011, doi: 10.1109/TPWRD.2010.2102778.

[9] J. Ma, Z. Wang, and Y. Liu, “Identifying transformer inrush current based on normalized grille curve,” *IEEE Transactions on Power Delivery*, vol. 26, no. 2, pp. 588-595, 2011, doi: 10.1109/TPWRD.2010.2101087.

[10] B. He, Z. Xuesong, and Q.B. Zhiqian, “A new method to identify inrush current based on error estimation,” *IEEE Transactions on power delivery*, vol. 21, no. 3, pp. 1163-1168, 2006, doi: 10.1109/TPWRD.2005.861337.

[11] R. Rahnavard, M. Valizadeh, A .A. B. Sharifian, and S. H. Hosseini, “Analytical analysis of transformer inrush current and some new techniques for its reduction,” *Proc. of Int. Conference on Power Systems and Transients*, 2005.

[12] Y. Mahmoudian, S. Sanati, A. Allameh, and A. R. Baghaei, "The Effect of Sympathetic Inrush Current on the Protection Maloperation in the Capacitor Feeder at Shahroud 63kV Substation," *15th International Conference on Protection and Automation of Power Systems (IPAPS),* Shiraz, Iran, 2020, pp. 80-83, doi: 10.1109/IPAPS52181.2020.9375609.

[13] M. A. Taghikhani, A. Sheikholeslami, and Z. Taghikhani, “Harmonic modeling of inrush current in core type power transformers using Hartley transform,” *Iranian Journal of Electrical and Electronic Engineering*, vol. 11, no. 2, pp. 174-183, 2015, doi: 10.22068/IJEEE.11.2.174.

[14] A. Alassi, K. Ahmed, A. Egea-Alvarez, and C. Foote, "Soft Transformer Energization: Ramping Time Estimation Method for Inrush Current Mitigation," *56th International Universities Power Engineering Conference (UPEC)*, Middlesbrough, United Kingdom, 2021, pp. 1-6, doi: 10.1109/UPEC50034.2021.9548203.

[15] Y. Pan, X. Yin, Z. Zhang, B. Liu, M. Wang, and X. Yin, “Three-phase transformer inrush current reduction strategy based on prefluxing and controlled switching,” *IEEE Access*, vol. 9, pp. 38961-38978, 2021, doi: 10.1109/ACCESS.2021.3062143.

[16] A. Moradi and S. M. Madani, “Technique for inrush current modelling of power transformers based on core saturation analysis,” *IET Generation, Transmission & Distribution*, vol. 12, no. 10, pp. 2317-2324, 2018, doi: 10.1049/iet-gtd.2017.1272.

[17] A. A. Nazari, F. Razavi, and A. Fakharian, "A novel method to differentiate internal faults and inrush current in power transformers using adaptive sampling and Hilbert transform," *Iranian Electric Industry Journal of Quality and Productivity*, vol. 11, no. 1, pp. 97-110, 2022.

[18] S. Bagheri, F. Safari, and N. Shahbazi, "Detection and Classification of Cross-Country Faults, Internal and External Electrical Faults and Inrush Current in Power Transformers Using Maximum Overlap Discrete Wavelet Transform," *Journal of Nonlinear Systems in Electrical Engineering*, vol. 8, no. 2, pp. 117-137, 2022.

[19] A. Yahiou, H. Mellah, and A. Bayadi, "Inrush Current Reduction by a Point-on-wave Energization Strategy and Sequential Phase Shifting in Three-Phase Transformer," *International Journal of Engineering*, vol. 35, no. 12, pp. 2321-2328, 2022, doi: 10.5829/IJE.2022.35.12C.07.

[20] S.- A. Hosseini *et al.*, "Proposing a New Approach to Generate the Differential Trajectory of the Differential Relays Using COMTRADE Files," *Sustainability*, vol. 14, no. 21, p. 13953, 2022, doi: 10.3390/su142113953.

[21] S. K. Gunda and V. S. S. S. S. Dhanikond, "Discrimination of Transformer Inrush Currents and Internal Fault Currents Using Extended Kalman Filter Algorithm (EKF)," *Energies*, vol. 14, no. 19, pp. 1-20, doi: 10.3390/en14196020.

[22] L. Wang *et al.,*"Simulation study on transformer inrush current and its suppression." *Journal of Physics: Conference Series*, vol. 2029, no. 1, 2021, doi: 10.1088/1742-6596/2029/1/012010.

[23] J. Mitra, X. Xu, and M. Benidris, "Reduction of Three-Phase Transformer Inrush Currents Using Controlled Switching," in *IEEE Transactions on Industry Applications*, vol. 56, no. 1, pp. 890-897, Jan.-Feb. 2020, doi: 10.1109/TIA.2019.2955627.

[24] G. de J Martínez-Figueroa, F. Córcoles-López, and S. Bogarra, "FPGA-Based Smart Sensor to Detect Current Transformer Saturation during Inrush Current Measurement," *Sensors*, vol. 23, no. 2, 2023, doi: 10.3390/s23020744.

[25] M. Changizian and A. Shoulaie, "Research on Effective Parameters on Overvoltage and Inrush Current During Starting-up of VSC-HVDC System," *Electromechanical Energy Conversion Systems*, vol. 1, no. 2, pp. 29-40, 2021, doi: 10.30503/eecs.2019.110965.

[26] C.W.T. McLyman, *Transformer and inductor design handbook*. CRC press, 2017.

[27] H. L. Chan, K. W. E. Cheng, T. K. Cheung, and C. K. Cheung, "Study on Magnetic Materials Used in Power Transformer and Inductor," *2nd International Conference on Power Electronics Systems and Applications*, Hong Kong, China, 2006, pp. 165-169, doi: 10.1109/PESA.2006.343091.

[28] R. Hilzinger and R. Werner, *Magnetic materials: fundamentals, products, properties, applications.* Vacuumschmelze, 2013.

[29] A.G. EPCOS, *Ferrites and accessories-SIFERRIT material N87,Data Sheet.* September 2006.

[30] Handbook, Ferroxcube Data, “Soft Ferrites and Accessories,” 2009.

Journal of Southern
Communication
Engineering

Islamic Azad University Bushehr Branch

Online ISSN: 2980-9231

Vol. 14, Issue 56, Summer 2025

Contents

Journal of Southern Communication Engineering
Islamic Azad University Bushehr Branch

Journal of Southern Communication Engineering

License holder: Islamic Azad University Bushehr Branch

Address: Iran, Bushehr, Islamic Azad University Bushehr Branch
Tel: (+98) 9107837420 **Fax:** (+98) 771 5683717
Website: https://jce.bushehr.iau.ir/ **eISSN:** 2980-9231
Email: jce.iaub@gmail.com – jce@ iaubusher.ac.ir

Indexed by:

JCE is published according to 87/422092 permission of Islamic Azad University.