Electronic & Cyber Defense
Vol. 14, No.1, 2026 (Serial No. 35) .pp. 1-20
ISSN: 2322-4347, E-ISSN: 2980-8979

Privacy-Preserving Task Matching Scheme Based on Similarity Search for Medical Consultation

Faezeh Nayyeri[1], Nasrollah Pakniat[2*], Ziba Eslami[3]
[1] PhD Student, Shahid Beheshti University, Tehran, Iran. *f_nayyeri@sbu.ac.ir*
[2] Assistant Professor, Iranian Research Institute for Information Science and Technology of Iran (IranDoc), Tehran, Iran. (*Correspondence: *Pakniat@irandoc.ac.ir*)
[3] Associate Professor, Shahid Beheshti University, Tehran, Iran. *z_eslami@sbu.ac.ir*

ARTICLE INFO

Article history:

Article Type: Research paper
Received: 03 January 2026
Revised: 23 February 2026
Accepted: 05 April 2026
Available online: 21 April 2026

Keywords:
Searchable Encryption
Task Matching
Privacy-Preserving
Similarity Search

ABSTRACT

With the growing expansion of remote healthcare services, the management and access to patient and physician information through electronic health records (EHRs) have become increasingly feasible. One of the most significant advantages of such systems is the ability to match patients with physicians based on patients' specific needs, which not only reduces waiting times but also improves the overall quality of medical services. However, privacy concerns related to the storage and transmission of sensitive medical data on servers pose serious challenges, including the risk of patient data leakage or exposure of physician information to malicious attacks. Consequently, the development of a secure and efficient matching system that ensures data confidentiality has become a critical necessity. To address these challenges, the use of cryptographic tools for protecting data prior to uploading it to cloud servers or other storage platforms has been proposed as an effective solution. Nevertheless, conventional encryption methods impose limitations on service optimization due to their inability to support search and operations over encrypted data. Therefore, this paper proposes a novel mechanism for secure and efficient patient-physician matching, enabling multi-keyword search over encrypted data in accordance with patient-specific requirements. The proposed scheme employs an innovative technique based on inner product computation to facilitate similarity measurement and matching between two vectors, which not only enhances efficiency but also improves the accuracy of the results. Security analysis confirms that the proposed scheme effectively preserves the privacy of both patient and physician data without disclosing any sensitive information. Compared to prior models with a time complexity of O (nmk), the proposed scheme significantly reduces the computational complexity to 0 (n) while achieving higher matching accuracy and more efficient search speed. Furthermore, by supporting weighted vectors, multi-keyword queries, and employing ASPE, it offers enhanced security against KPA and CPA attacks.

Cite this article: Nayyeri, F., Pakniat, N., Eslami, Z. (2026). Privacy-Preserving Task Matching Scheme Based on Similarity Search for Medical Consultation. *Electronic and Cyber Defense*,14(1),1-20.2026.

DOI.https://dor.isc.ac/dor/20.1001.1.23224347.1405.14.1.1.3

Publisher: Imam Hossein University.

«پدافند الکترونیک و سایبری»
سال چهاردهم، شماره ۱، بهار ۱۴۰۵؛ ص ۹۹-۱
شاپا الکترونیکی: ۸۹۷۹-۲۹۸۰ شاپا چاپی: ۴۳۴۷-۲۳۲۲

طرح تطبیق وظایف با حفظ حریم خصوصی کاربران بر اساس جست‌وجوی شباهت جهت مشاوره پزشکی

فائزه نیری[۱]، نصراله پاک‌نیت[۲*]، زیبا اسلامی[۳]

[۱] دانشجوی دکتری، دانشگاه شهید بهشتی، تهران، ایران. *f_nayyeri@sbu.ac.ir*

[۲] استادیار، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)، تهران، ایران (نویسنده مسئول: *Pakniat@irandoc.ac.ir*)

[۳] دانشیار، دانشگاه شهید بهشتی، تهران، ایران. *z_eslami@sbu.ac.ir*

چکیده

با گسترش روزافزون خدمات درمانی غیرحضوری، مدیریت و دسترسی به اطلاعات پزشکان و بیماران از طریق پرونده‌های الکترونیکی سلامت فراهم شده است. یکی از مهم‌ترین مزایای این سیستم‌ها، امکان تطابق بین بیماران و پزشکان بر اساس نیازهای خاص بیماران است که علاوه بر کاهش زمان انتظار، به بهبود کیفیت خدمات درمانی نیز کمک می‌کند. بااین‌حال، چالش‌های مرتبط با حفظ حریم خصوصی بیماران و پزشکان در فرایند ذخیره‌سازی و انتقال داده پزشکی روی سرورها، متأثر از این نگرانی است که اطلاعات حساس بیماران درز پیدا کرده و یا اطلاعات پزشکان دستخوش حملات مهاجمان قرار گیرد؛ بنابراین ایجاد سیستم تطابق کارآمدی که امنیت داده‌ها را تضمین کند، به یک نیاز اساسی تبدیل شده است. به‌منظور مقابله با این چالش‌ها، استفاده از ابزارهای رمزنگاری باهدف محافظت از داده‌ها قبل از بارگذاری آن‌ها روی سرورهای ابری یا سایر سیستم‌های ذخیره‌سازی، به‌عنوان راهکار مؤثری پیشنهاد شده است. باوجوداین، روش‌های سنتی رمزنگاری به دلیل عدم امکان جستوجو و انجام عملیات روی داده‌های رمزگذاری شده، محدودیت‌هایی در ارائه خدمات بهینه ایجاد می‌کنند؛ بنابراین، در این مقاله، مکانیزم نوینی برای ایجاد تطابق امن و کارآمد بین بیماران و پزشکان ارائه‌شده است که با درنظرگرفتن نیازهای بیمار، امکان جستوجوی چند کلیدواژه روی داده‌های رمزگذاری شده را فراهم می‌کند. در طرح پیشنهادی از روشی نوین بر پایه محاسبه ضرب داخلی برای تطابق و اندازه‌گیری شباهت بین دو بردار استفاده شده است که علاوه بر بهبود کارایی، دقت نتایج حاصلشده را نیز افزایش می‌دهد. تجزیه‌وتحلیل‌های امنیتی انجام‌شده نشان می‌دهند که طرح ارائه‌شده از حریم خصوصی داده‌های پزشکان و بیماران به طور مؤثری محافظت می‌کند و هیچ اطلاعاتی را فاش نمی‌سازد. در مقایسه با مدل‌های پیشین با پیچیدگی زمانی $O(nmk)$، مدل پیشنهادی با کاهش چشمگیر پیچیدگی به $O(n)$، دقت تطبیق بالاتر و سرعت جستوجوی بهینه‌تری را ارائه می‌دهد. همچنین، با پشتیبانی از بردارهای وزن‌دار، چند کلیدواژه و بهره‌گیری از ASPE، سطح امنیتی بالاتری در برابر حملات KPA و CPA فراهم می‌سازد.

مشخصات مقاله

تاریخچه مقاله:

نوع مقاله: علمی پژوهشی

دریافت: ۱۴۰۴/۱۰/۱۳

بازنگری: ۱۴۰۴/۱۲/۰۴

پذیرش: ۱۴۰۵/۰۱/۱۶

ارائه آنلاین: ۱۴۰۵/۰۲/۰۱

کلید واژه ها:

رمزگذاری قابل جستوجو

تطبیق وظایف

حفظ حریم خصوصی

جست‌وجوی شباهت

استناد: نیری، فائزه، پاک‌نیت، نصراله، اسلامی، زیبا. (۱۴۰۵). طرح تطبیق وظایف با حفظ حریم خصوصی کاربران بر اساس جست‌وجوی شباهت جهت مشاوره پزشکی. پدافند الکترونیکی و سایبری، ۱۴(۱)، ۲۰-۱. DOR: https://dor.isc.ac/dor/20.1001.1.23224347.1405.14.1.1.3

ناشر: دانشگاه جامع امام حسین(ع).

۱. مقدمه

با گسترش جوامع بشری و همه‌گیری بیماری‌های اپیدمی، تقاضای فزاینده‌ای برای دریافت خدمات پزشکی غیرحضوری به وجود آمده است و پزشکان و متخصصان ناچار به انتقال فعالیت‌های خود به فضای مجازی شده‌اند. البته پیشرفت‌های فناورانه و به‌وجودآمدن ابزارهای نرم‌افزاری، تأثیر بسیاری در بهبود عملکرد مدیریت قرارهای ملاقات، تشخیص، پیشنهاد مناسب‌ترین پزشک، ذخیره‌سازی، تضمین دردسترس‌بودن خدمات سلامتی و در نهایت کاهش هزینه‌ها داشته است [۱]. این ابزارها، با ارائه امکان دریافت خدمت غیرحضوری به بیماران، امکان اختصاص زمانی بیشتر به درمان را فراهم کرده و در نهایت به بهبود بهره‌وری و تجربهٔ کلی پزشکان و بیماران کمک می‌کنند [۲].

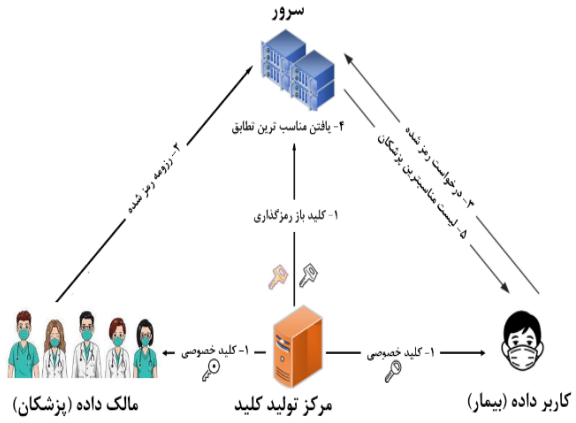
یکی از ابزارهای استفاده‌شده برای بارگذاری اطلاعات پزشکان و درخواست بیماران به‌منظور یافتن مناسب‌ترین پزشک باتوجه‌به نیازمندی آن‌ها، سیستم‌های جمع‌سپاری است که این امکان را در اختیار کاربران قرار می‌دهد تا داده‌های خود را برای ذخیره‌سازی، مدیریت و ارائه به سرورهای بیرونی بسپارند [۳]. سیستم‌های جمع‌سپاری در سال‌های اخیر، روی انواع مختلف پلتفرم‌ها توسعه یافته‌اند. در این سیستم‌ها و به طور خاص در حوزه پژوهش این مقاله (مشاوره پزشکی)، ازیک‌طرف تمامی پزشکان اطلاعات مربوط به رزومه کاری خود و از طرف دیگر هر بیمار نیازمندی‌های خود را روی سرور بارگذاری کرده و سرور مسئولیت تطابق بین این دو گروه را بر اساس مشخصات اعلام شده بر عهده دارد. در این سناریو، سرور به دنبال یافتن مناسب‌ترین پزشکان برای معرفی به بیمار برای مشاوره است؛ به‌طوری که بیمار بدون مراجعهٔ حضوری بتواند از خدمات او بهره‌مند شود.

یکی از مسائل مهم در سناریوهای جمع‌سپاری، امنیت و حفاظت از حریم خصوصی کاربران است. در این راستا، داده‌ها قبل از بارگذاری روی سرور رمزگذاری می‌شوند [۴]. زمانی که داده‌ها توسط طرح‌های رمزگذاری سنتی رمزگذاری می‌شوند، انجام محاسبات و جست‌وجو روی آن‌ها غیرممکن می‌شود. در این حالت، کاربر می‌تواند ابتدا تمام متون رمز شده مربوطه را دریافت (دانلود) و سپس با استفاده از کلید خصوصی، داده‌های اصلی را برای محاسبات موردنظر بازیابی کند. این فعالیت هزینه دریافت داده‌ها، رمزگشایی آن‌ها و سربار ذخیره‌سازی زیادی را برای کاربران به همراه خواهد داشت [۵]. برای حل این مسئله، می توان از روش‌های رمزگذاری قابل جست‌وجو [۶] استفاده کرد. این روش‌ها ضمن حفظ محرمانگی داده‌های رمز شده، به کاربران امکان جست‌وجوی امن روی داده‌ها را می‌دهند [۷]. بااین‌حال، به دلیل حجم زیاد کاربران و وظایف در سیستم‌های جمع‌سپاری، روش‌های رمزگذاری جست‌وجوی سنتی برای این منظور مناسب نیست و از کارایی کمتری رنج می‌برند [۸].

مفهوم تطبیق امن وظایف باهدف غلبه بر این چالش، ضمن درنظرگرفتن تعداد روزافزون کاربران در پلتفرم‌های جمع‌سپاری ایجاد شده است. بااین‌حال، بررسی پیشینه مسئله بیانگر این است که تاکنون طرح تطبیق وظایف امنی به‌منظور ارائه خدمات مشاورهٔ پزشکی ارائه نشده است. باتوجه‌به این مهم، در این مقاله، در ابتدا به مدل‌سازی مسئله پرداخته شده و در ادامه، طرحی تطبیق وظایف امن و کارا برای اختصاص مناسب‌ترین پزشکان به بیماران ارائه شده است. در طرح پیشنهادی، از ضرب داخلی [۱] (ضرب اسکالر [۲]) به‌عنوان معیار تشابه برای تعیین مناسب‌ترین تطابق استفاده می‌شود و با انجام برخی پیش‌پردازش‌ها روی داده‌ها، دقت و کارایی بیشتری ایجاد می‌شود. نوآوری‌های طرح پیشنهادی در مقایسه با طرح‌های موجود به شرح زیر است:

- مدل‌سازی و ارائه اولین طرح تطبیق وظایف امن با کاربرد در حوزه ارائه خدمات مشاوره پزشکی و اختصاص مناسب‌ترین پزشکان به بیماران.
- افزایش دقت و کارایی با به‌کارگیری نمایشی جدید برای بردارهای متناظر با پزشکان و بیماران.
- فراهم‌کردن امکان جست‌وجوی چند کلیدواژه‌ای روی داده‌های رمزنگاری شده که منجر به بهبود تجربه پزشکان و افزایش رضایتمندی بیماران می‌گردد.

همان‌طور که در شکل (۱) مشخص است، ۴ موجودیت مرکز تولید کلید [۳]، سرور [۴]، مالک داده [۵] و کاربر داده [۶] در مدل سیستم پیشنهادی این مقاله وجود دارند که در ادامه وظایف هر یک تشریح شده است:



شکل (۱): مدل سیستم رمزگذاری قابل‌جست‌وجو استفاده شده

[1] Inner Product.
[2] Scalaer Product.
[3] Key Generation Center.
[4] Server.
[5] Data Owner.
[6] Data User.

مرکز تولید کلید: این موجودیت، کلیدهای خصوصی و کلیدهای باز رمزگذاری[1] را تولید کرده و آن‌ها را در اختیار دیگر موجودیت‌های سیستم قرار می‌دهد.

سرور: انبار اطلاعاتی بزرگی است که توانایی محاسباتی زیادی دارد و اطلاعات پزشکان را در خود نگه می‌دارد. علاوه‌برآن، سرور وظیفه یافتن و تخصیص مناسب‌ترین پزشکان به بیماران را بر عهده دارد.

مالک داده: موجودیتی است که داده‌ها را تولید و رمزگذاری کرده و سپس آن‌ها را روی سرور بارگذاری می‌کند. مالک داده، داده‌ها را به‌گونه‌ای رمزگذاری می‌کند که قابلیت جست‌وجو روی آن‌ها وجود داشته باشد. در این مقاله منظور از مالکان داده، پزشکانی هستند که با استفاده از کلید خصوصی، به تولید متون رمز شده یا نمایه[2] پرداخته و آن‌ها را به همراه اطلاعات مربوطه که رمز شده‌اند، روی سرور بارگذاری می‌کنند.

کاربر داده: این موجودیت پرس‌وجوهای رمزگذاری شده را که دریچه[3] نامیده می‌شوند، برای جست‌وجوی داده‌ای مشخص به سرور ارسال می‌کند. در این مقاله منظور از کاربران داده بیمارانی هستند که با استفاده از کلید خصوصی خود، نیازمندی‌هایشان را برای یافتن پزشک مناسب برای مشاوره غیرحضوری، رمزگذاری و به سرور ارسال می‌کنند.

در ادامه، ساختار مقاله به‌صورت زیر است: در بخش ۲، پیشینهٔ پژوهش و مطالعات مرتبط انجام‌شده در این حوزه بررسی می‌شود. در بخش ۳، مسئله اصلی پژوهش و چالش‌های موجود به طور دقیق تبیین می‌شود. بخش ۴ مربوط به پیش‌نیاز استفاده‌شده در مدل پیشنهادی است. در بخش ۵ و ۶ به‌تفصیل روش پیشنهادی و درستی آن شرح داده شده است. در بخش ۷، تحلیل امنیتی مدل پیشنهادی انجام‌شده و مقاومت آن در برابر انواع حملات احتمالی ارزیابی می‌شود. مدل داده‌ها، پیاده‌سازی، نتایج و پیچیدگی محاسباتی مدل و کارایی آن از نظر زمان اجرا در بخش ۸ آورده شده است. در نهایت، در بخش ۹، نتایج به‌دست‌آمده از این پژوهش به همراه پیشنهادهایی برای تحقیقات آتی در این حوزه ارائه می‌شود.

۲. پیشینه تحقیق

سال‌هاست که سیستم‌های جمع‌سپاری، حوزه‌های مختلف را تسخیر کرده است و سازمان‌ها برای تجارت و همچنین طراحی پلتفرم‌های برخط از این سیستم‌ها استفاده می‌کنند [۳]. به همین دلیل، در سال‌های اخیر مطالعات متعددی در زمینه تطبیق وظایف در سیستم‌های جمع‌سپاری، انجام شده است [۴، ۵، ۹].

سیستم‌های جمع‌سپاری بر اساس نحوه تطابق وظایف به دودسته «وظایف منتخب عامل» [4] (WST) و «وظایف محول شده توسط سرور» [5] (SAT) تقسیم می‌شوند [۶]. در WST سرور وظایف را به‌صورت برخط از طرف مالکان داده منتشر کرده و عامل‌ها مطابق نیازمندی‌های خود و به طور مستقل، وظایف را انتخاب می‌کنند. این در حالی است که در SAT، سرور خود مسئول تخصیص وظایف به عامل‌ها است. مزیت دسته دوم در این است که سرور می‌تواند تمامی عامل‌ها را در زمان تخصیص وظایف در نظر گیرد. به‌این‌ترتیب، هر عامل به طور مستقل وظایف را انتخاب نکرده و از این طریق، احتمال باقی‌ماندن برخی از وظایف بدون عامل و یا وجود چندین عامل برای یک وظیفه کاهش می‌یابد.

بااین‌حال، ارسال اطلاعات کاربران به سرور به‌صورت خام، تهدیدی برای حفظ حریم خصوصی آن‌هاست [۷] که برای حل این مشکل از روش‌های رمزگذاری استفاده می‌شود [۵، ۸، ۱۰]. در این رویکردها، داده‌ها قبل از بارگذاری روی سرور، باید رمزگذاری شوند. در برخی از روش‌های ارائه شده [۱۱]، از رمزگذاری تنها برای حفظ محرمانگی اطلاعات یکی از موجودیت‌های مالک داده یا کاربر داده استفاده شده است. بااین‌حال، در سال‌های اخیر مسئله حفظ حریم خصوصی تمامی کاربران (مالکان داده و کاربران داده) در سیستم‌های جمع‌سپاری به طور گسترده موردتوجه قرار گرفته است [۱۲، ۱۳].

از طرف دیگر، به دلیل عدم اعتماد کامل به سرور، برخی روش‌های پیشنهادشده برای اجرای تطبیق وظایف روی داده‌های رمزگذاری شده، به یک مرجع قابل‌اعتماد متکی هستند [۱۴]. بااین‌حال، سرور در این روش‌ها قادر است اطلاعات کاربران را بر اساس نتایج تطبیق، استخراج کند. باتوجه‌به عدم امکان اعتماد کامل به سرورها، استفاده از این روش‌ها در صورت حساسیت و محرمانگی زیاد داده‌های کاربران، نامناسب است. مسئله مهم دیگری که وجود دارد این است که وقتی داده‌ها رمزگذاری می‌شوند، توانایی سرور برای جست‌وجوی کلیدواژه‌ها[6] محدود می‌شود و سرور به‌سادگی نمی‌تواند کلیدواژه متن اصلی[7] را روی داده‌های رمز شده جست‌وجو کند. به‌منظور حفظ محرمانگی داده‌ها، به روش‌هایی باقابلیت جست‌وجوی کلیدواژه روی داده‌های رمز شده نیاز است که هیچ‌گونه اطلاعاتی درباره کلمات جست‌وجو شده و یا اسناد بازیابی شده افشا نشود [۱۵، ۱۶]. جست‌وجو روی داده‌های رمز شده در بسیاری از پژوهش‌ها موردتوجه قرار گرفته و روش‌های مختلفی برای این منظور پیشنهاد شده است. طرح‌های سنتی رمزگذاری قابل جست‌وجو [۱۲، ۱۷] صرفاً برای سناریوهای تک‌کاربره[8] قابل‌استفاده هستند، بدین معنی که در این طرح‌ها، تنها دارنده کلید مخفی می‌تواند روی داده‌های رمز شده

[1] Re-Encryption Key
[2] Index
[3] Trapdoor

[4] Worker Selected Tasks
[5] Server Assigned Tasks
[6] Keyword Search
[7] Plaintext
[8] Single-User

جست‌وجو کند. برای حل این مشکل، رمزگذاری قابل جست‌وجوی چندکاربره[1] ارائه شده است. در این نوع از رمزگذاری قابل جست‌وجو، مالک و هرکدام از کاربران مجاز، کلید خصوصی منحصربه‌فرد خود را داشته و مالک داده می‌تواند داده‌ها را به‌گونه‌ای رمزگذاری کند که هرکدام از کاربران مجاز بتوانند با استفاده از کلید خصوصی خود، به جست‌وجو روی اطلاعات رمزگذاری شده بپردازند [۱۱].

شو و همکارانش [۷] روش رمزگذاری قابل جست‌وجوی چندکاربره با تطبیق تک کلیدواژه را ارائه کردند که علاوه بر محافظت هم‌زمان از حریم خصوصی کاربران، قابلیت جست‌وجوی کلیدواژه را نیز فراهم می‌کرد. با پیشرفت تحقیقات، تطبیق چند کلیدواژه [۱۲] و ناشناس بودن کاربران [۱۸] نیز توسعه یافت که باتوجه‌به استفاده از روش باز رمزگذاری نیابتی[2]، هیچ‌یک از آن‌ها نمی‌توانستند از نشت کلیدهای باز رمزگذاری جلوگیری کنند.

پوپا و همکارانش [۱۹] مفهوم رمزگذاری قابل جست‌وجوی چند کلیدی[3] را پیشنهاد کردند که در آن کاربران می‌توانند اسناد رمزگذاری شده با کلیدهای مختلف را جست‌وجو کنند. باوجوداین، تعداد کمی از طرح‌ها می‌توانند مسئله تطبیق را به طور هم‌زمان بر اساس ویژگی‌های مختلف پشتیبانی کنند [۲۰].

علاوه بر طرح‌های متداول فوق، سایر روش‌ها در حوزه رمزگذاری قابل جست‌وجو مانند «جست‌وجوی کلیدواژه مبتنی بر پخش»[4] [۲۱]، «جست‌وجوی کلیدواژه مبتنی بر ویژگی»[5] [۲۲] و «جست‌وجوی کلیدواژه مبتنی بر هویت»[6] [۲۳] معرفی شده‌اند که بر اساس مجوز مالک داده‌اند. چنین رویکردهایی برای سیستم‌های جمع‌سپاری که جست‌وجو به‌صورت آزادانه صورت می‌پذیرد، جایگاهی ندارند.

اجرای عملگر جست‌وجو برای تمام کاربران در پایگاه‌داده مربوطه منجر به پیچیدگی محاسباتی می‌شود که به‌صورت خطی با تعداد کاربران مقیاس می‌شود [۲۴]. برای پرداختن به این موضوع و افزایش کارایی، فن‌های یادگیری ماشینی به‌عنوان راه‌حل استفاده شده است [۲۵, ۲۶]. لیکن استفاده‌کردن از دسته‌بندی و خوشه‌بندی در مسئله پیش‌رو باتوجه‌به ماهیت آن، ممکن است به حملات استنتاجی منجر شده و سرور را قادر سازد تا با استفاده از نتایج تطابق، اطلاعات کاربران را استنتاج کند. در همین راستا، طرح محاسباتی ایمن k-نزدیک‌ترین همسایه[7] (k-NN) برای پایگاه‌های داده رمزگذاری شده با استفاده از مفهوم «رمزگذاری نامتقارن حفظ‌کننده ضرب اسکالر»[8] (ASPE) معرفی شده است [۲۷]. از لحاظ عملکردی، ASPE رفتاری مشابه ضرب داخلی ایمن[9] دارد که با روش‌های متفاوت پیاده‌سازی می‌شود. مزیت کلیدی ASPE حفظ ارزش‌های ضرب داخلی است که آن را به‌ویژه برای جست‌وجوهایی با مرتبه زیرخطی بر اساس پرس‌وجوهای شباهت مناسب می‌کند. پیشرفت‌های اخیر، کاربرد ASPE را در رایانش ابری بیشتر افزایش داده است [۲۸].

هدف این مقاله ارائه یک طرح تطابق امن و کارا برای به‌کارگیری در حوزه مشاوره پزشکی است که باتوجه‌به کاربرد در نظر گرفته شده، پزشکان به‌عنوان مالکان داده و بیماران به‌عنوان کاربران داده هستند و لازم است از حریم خصوصی هر کاربر محافظت شود. در این راستا در جدول زیر، مقایسه‌ای بین مدل پیشنهادشده با طرح‌های مرتبط قبلی از نظر عملکرد و امنیت انجام شده است.

جدول (۱): مقایسه عملکرد و امنیت مدل پیشنهادی با طرح‌های مرتبط

چند کلیدواژه	بردار وزن‌دار	استفاده از ASPE	پیچیدگی زمانی	مقاوم در برابر حملات	مقاله
✓	✓	×	O(nmk)	KCM و KBM	ML-RKS [25]
✓	×	×	O(n log n)	KPM	[12]
✓	×	✓	O(n)	KPA و CPA	[29]
✓	✓	✓	O(n)	KPA و CPA	مدل پیشنهادی

۳. بیان مسئله

همان‌طور که بیان شد، مسئله موردنظر در این مقاله دارای ۴ نوع موجودیت مرکز تولید کلید، سرور، مالک داده و کاربر داده است. مالک داده و کاربر داده به ترتیب با عناوین درخواست‌کننده[10] و پاسخ‌دهنده[11] نیز شناخته می‌شوند. هر دو نوع کاربر، توانمندی‌ها و شرایط خود را روی سرور بارگذاری می‌کنند و سرور مسئول اختصاص مناسب‌ترین درخواست‌کنندگان به پاسخ‌دهندگان است. این تطابق معمولاً باتوجه‌به شرایط دو طرف صورت می‌پذیرد. فرض این مقاله بر این است که مرکز تولید کلید، یک موجودیت کاملاً قابل‌اعتماد[12] بوده و ارتباط بین این مرکز و سایر موجودیت‌ها نیز امن است. این در حالی است که سرور موجودیتی نیمه صادق[13] در نظر گرفته می‌شود؛ به این معنی که

[1] Multi-User
[2] Proxy Re-Encryption
[3] Multi-key Searchable Encryption
[4] Broadcast-Based Keyword Search
[5] Attribute-Based Keyword Search
[6] Identity-Based Keyword Search
[7] K-Nearest Neighbors
[8] Asymmetric Scalar-Product-Preserving Encryption (ASPE)
[9] Secure Inner Product
[10] Requester
[11] Worker
[12] Fully Trusted
[13] Semi-Honest

در عین انجام وظایف و رعایت قوانین و روندها، به دنبال یافتن ارتباط بین داده‌ها و کلیدهای خصوصی است.

در مدل در نظر گرفته‌شده هیچ‌گونه تبانی بین موجودیت‌ها وجود ندارد؛ ازاین‌رو، مدل ارائه شده باید تضمین کند که سرور قادر به استنتاج اطلاعات حساس نمایه‌ها، دریچه‌ها و ارتباط بین آن‌ها از روی متون رمز شده متناظرشان نیست. علاوه بر این، مدل پیشنهادی در عین حفظ محرمانگی الزامات پزشکان و بیماران، باید در مقابل حملات زیر مقاوم باشد:

- **حمله با متن اصلی معلوم**[1] (KPA): حمله‌ای که با استفاده از تعدادی متن اصلی و متن رمز شده[2] متناظر با آن‌ها انجام می‌شود. در این نوع حمله، سرور می‌تواند چندین جفت از متن اصلی و متن رمز شده را به دست آورد و با استفاده از اطلاعات این جفت‌ها، سعی کند اطلاعاتی درباره دیگر متون رمز شده یا کلیدهای مخفی به دست آورد.
- **حمله با متن اصلی منتخب**[3] (CPA): حمله‌ای که در آن مهاجم توانایی انتخاب متن‌های اصلی و متن‌های رمز شده متناظر را دارد و با استفاده از اطلاعات این جفت‌ها، سعی می‌کند اطلاعاتی درباره دیگر متون رمز شده و یا کلیدهای مخفی به دست آورد. به بیان دیگر، در این نوع حمله سرور می‌تواند متون رمز شده متناظر با چندین متن اصلی را به انتخاب خود مشاهده کند. مشخص است که در این نوع حمله، مهاجم قدرت بیشتری نسبت به مدل KPA دارد.

به طور مشخص، اهداف زیر در طرح تطبیق وظایف ارائه‌شده محقق می‌گردد:

- **جست‌وجوی چند کلیدواژه:** طرح ارائه‌شده، از جست‌وجوی چند کلیدواژه در تطبیق وظایف پشتیبانی می‌کند.
- **جست‌وجوی چندکاربره:** طرح ارائه‌شده، از جست‌وجوی چند درخواست‌کننده برای تطبیق وظایف پشتیبانی می‌کند.
- **محرمانگی نمایه و محرمانگی دریچه**: روش ارائه‌شده تضمین می‌کند که سرور نمی‌تواند اطلاعات حساس نمایه‌ها و دریچه‌ها را استخراج کند.
- **عدم پیوند نمایه و عدم پیوند دریچه**: در فرایند تطبیق وظایف، نمایه‌ها و دریچه‌ها در اختیار سرور قرار می‌گیرند. به‌منظور محافظت از حریم خصوصی، نمایه‌ها و دریچه‌ها به‌گونه‌ای هستند که سرور نتواند رابطه میان آن‌ها را با استفاده از متن‌های رمزی متناظر آن‌ها استنتاج کند. این بدین معنی است که سرور نمی‌تواند تشخیص دهد که آیا دو پاسخ‌دهنده یا دو درخواست‌کننده دارای کلیدواژه‌های یکسان هستند یا خیر.

[1] Known Plaintext Attack (KPA)
[2] Ciphertext
[3] Chosen Plaintext Attach (CPA)

علاوه‌برآن، متون رمز شده موردنیاز یکسان در فرایندهای تطبیق وظایف مختلف، متفاوت هستند.

نمادهای استفاده شده در این بخش در جدول (۲) آورده شده است.

در مدل پیشنهادی، به‌منظور نمایش اطلاعات مربوط به پزشکان و درخواست بیماران از مدل فضای برداری[4] استفاده شده است [۳۰]. در این راستا، برای مدل‌کردن اطلاعات هر پزشک از دو بردار استفاده می‌شود. یک بردار دودویی n-تایی که بردار "نوع خدمت" نامیده و با W نشان داده می‌شود که نوع تخصص و فوق‌تخصص پزشک را نشان می‌دهد و بیانگر نوع خدمتی است که یک پزشک می‌تواند ارائه دهد؛ و یک بردار ۱۰-تایی که بردار "امتیاز و زمان انتظار" نامیده و با V نشان داده می‌شود که نیمی از آن بیانگر میزان رضایت بیماران از پزشک و نیمی دیگر زمان انتظار تقریبی برای دریافت مشاوره را نشان می‌دهد. باتوجه‌به توضیحات فوق، می‌توان پایگاه‌داده m پزشک را به‌صورت $D = \{d_1, d_2, \dots, d_m\}$ نشان داد که هر شیء از این پایگاه‌داده به‌صورت $d_i = \{W_i, V_i\}$ است.

جدول (۲): نمادهای استفاده شده به همراه توضیحات

نماد	توضیحات
W	بردار نوع خدمت
n	طول بردار نوع خدمت
V	بردار امتیاز و زمان انتظار
n_1	طول بردار امتیاز
n_2	طول بردار زمان انتظار
D	پایگاه‌داده پزشکان
d_i	پزشک i-ام از پایگاه‌داده
m	تعداد پزشکان
q	بردار درخواست بیمار
k	تعداد پزشکان برگردانده شده به‌عنوان نتیجه

بدون ازدست‌رفتن کلیت مسئله فرض می‌شود که پزشکان در سه دسته عمومی، متخصص و فوق‌تخصص تقسیم‌بندی شده‌اند.

[4] Vector Space Model (VSM)

همچنین فرض می‌شود تعداد انواع تخصص و فوق‌تخصص در سیستم به ترتیب برابر با n_1 و n_2 باشد به‌طوری که $n_1 + n_2 = n - 1$. در این صورت با افزودن یک مؤلفه متناظر برای پزشکان عمومی به مجموعه تخصص‌ها و فوق‌تخصص‌ها، مجموعه n-تایی {عمومی، تخصص ۱، و... تخصص n_1، فوق‌تخصص ۱، و...، فوق‌تخصص n_2} ایجاد می‌شود که می‌توان با استفاده از آن، هر پزشک را با یک بردار دودویی n-تایی نشان داد. در بردار موردنظر، درصورتی‌که پزشک عمومی باشد تنها اولین درایه که مربوط به پزشک عمومی است، برابر ۱ خواهد بود و باقی درایه‌ها برابر ۰ خواهند بود. درصورتی‌که پزشک متخصص یا فوق‌تخصص باشد، درایه‌های متناظر با تخصص و فوق‌تخصص او ۱ و سایر درایه‌های آن ۰ خواهد بود. علاوه‌برآن، فرض بر این است که امتیاز پزشک عددی صحیح در بازه ۱ تا ۵ باشد که این مقدار باتوجه‌به ارزیابی‌های بیماران پیشین از پزشک و ویژگی‌هایی مانند دوره‌های تخصصی گذرانده شده و دیگر ویژگی‌ها در بازه‌های زمانی مشخص محاسبه می‌شود که در آن ۱ پایین‌ترین امتیاز و ۵ بالاترین امتیاز است. زمان انتظار نیز عددی صحیح بین ۱ تا ۵ است که تعابیر آن در جدول زیر ارائه شده است:

جدول (۳): تعابیر اعداد برای زمان انتظار

عدد	توضیح T_i
۱	به معنی تا ۲۴ ساعت
۲	به معنی از ۱ تا ۳ روز
۳	به معنی از ۳ تا ۷ روز
۴	به معنی از ۷ تا ۱۵ روز
۵	به معنی از ۱۵ تا ۳۰ روز

در نمایش اولیه در نظر گرفته شده، هر یک از بردارهای "امتیاز" و "زمان انتظار" مربوط به هر پزشک، ۵ مؤلفه دارد که درایه متناظر با عدد متناظر با پزشک برابر با ۱ و سایر درایه‌ها برابر با صفر قرار داده می‌شود.

به‌منظور مدل‌کردن درخواست‌های بیماران نیز از بردارهایی مشابه با بردار پزشکان استفاده می‌شود؛ با این تفاوت که این بردار بر اساس نیازمندی بیمار تکمیل می‌گردد. هر درخواست از سمت بیمار به‌صورت $q_j = \{W_j, V_j\}$ نمایش داده می‌شود. هر بیمار، متناسب با درخواست خود می‌تواند پزشک عمومی یا متخصص و یا فوق‌تخصص انتخاب نماید. وی این کار را با ایجاد یک بردار دودویی n-تایی و برابر با ۱ قراردادن مولفه(های) متناظر با نیاز خود انجام می‌دهد. همچنین حداقل امتیاز پزشک درخواستی و حداکثر زمان انتظار که می‌تواند جهت مشاوره صبر نماید را تعیین می‌نماید. تعبیر این مقادیر مشابه قسمت پزشکان است.

به‌منظور شفاف‌شدن مدل‌سازی، یک مثال در شکل‌های (۲) و (۳) آورده شده است.

ردیف	عمومی	تخصص				فوق تخصص										امتیاز					زمان انتظار				
		a	b	c	d	a1	a2	b1	b2	c1	c2	d1	d2	۱	۲	۳	۴	۵	۱	۲	۳	۴	۵
۱	۱	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱	۰	۱	۰	۰	۰	۰
۲	۱	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱	۰	۰	۰	۰	۱	۰	۰	۰
۳	۰	۰	۱	۰	۰	۰	۰	۱	۰	۰	۰	۰	۰	۰	۰	۱	۰	۰	۰	۰	۱	۰	۰
۴	۰	۰	۱	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱	۰	۰	۰	۰	۰	۱	۰	۰
۵	۰	۰	۰	۰	۱	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱	۰	۰	۰	۰	۱	۰	۰	۰
۶	۱	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱	۰	۰	۱	۱	۰	۰	۰	۰	۰
۷	۱	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱	۰	۰	۰	۰	۰	۱	۰	۰
۸	۱	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱	۰	۱	۰	۱	۰	۰	۰	۰

شکل (۳): پایگاه داده مربوط به پزشکان. هر ردیف متناظر با یک پزشک است

درخواست	عمومی	تخصص				فوق تخصص										امتیاز					زمان انتظار				
بیمار		a	b	c	d	a1	a2	b1	b2	c1	c2	d1	d2	۱	۲	۳	۴	۵	۱	۲	۳	۴	۵
	۱	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱	۰	۱	۰	۰	۰	۰

شکل (۲): بردار مربوط به درخواست بیمار

در مثال آورده شده، بیمار به دنبال پزشک عمومی برای دریافت مشاوره است که حداقل امتیاز وی ۴ باشد و حداکثر طی ۱ تا ۳ روز آینده (مقدار ۲) بتواند از مشاوره وی بهره‌مند شود. در گام نخست، همان‌طور که در پایگاه‌داده پزشکان دیده می‌شود، پزشکان شماره ۱، ۲، ۶، ۷ و ۸ عمومی هستند و به‌عنوان کاندید اولیه انتخاب می‌شوند. در گام بعد، امتیاز و زمان انتظار آن‌ها مورد بررسی قرار می‌گیرد. باتوجه‌به اطلاعات موجود، از بین پزشکان عمومی، امتیاز پزشک شماره ۲ برابر ۳ (کمتر از مقدار تعیین شده توسط بیمار)، امتیاز پزشکان شماره ۷ و ۸ برابر ۴ (برابر با امتیاز موردنظر بیمار) و امتیاز پزشکان شماره ۱ و ۶ برابر ۵ (بیشتر از امتیاز موردنظر بیمار) است. در ادامه، زمان انتظار پزشکان شماره ۱، ۶ و ۸ کمتر یا مساوی زمان انتظار تعیین شده توسط بیمار است؛ بنابراین، سه پزشک مذکور به‌عنوان پزشکان مناسب، به بیمار معرفی می‌شوند. لازم به ذکر است در پژوهش پیش‌رو فرض بر این است که حداقل k پزشک با ویژگی‌های موردنیاز بیمار وجود دارد و درصورتی‌که تعدادی بیشتر از پزشکان حائز شرایط باشند، تعداد k تا از آن‌ها به‌عنوان پزشک مناسب توسط سرور انتخاب شده و به بیمار معرفی می‌شوند.

۴. پیش‌نیازها

- **مفهوم** ASPE

یکی از ابزارهای رمزنگاری که امکان محاسبه ایمن ضرب داخلی بین دو بردار را فراهم می‌کند، مفهوم ASPE است. این مفهوم به طور خاص در کاربردهای حفظ حریم خصوصی، مانند محاسبات امن چندجانبه[1]، جست‌وجوی رمزگذاری شده و داده‌کاوی با

[1] Secure Multi-Party Computing (SMPC)

حفظ حریم خصوصی مفید است. در ASPE، سرور قادر است که بدون دانستن مقادیر متن اصلی بردارهای W_1، W_2 و Q، حاصل‌ضرب داخلی آن‌ها را به‌گونه‌ای محاسبه نماید که ترتیب حاصل‌ضرب داخلی متون اصلی (به‌عنوان یک رابطه مقایسه‌ای) در حاصل‌ضرب متون رمز شده حفظ گردد. یعنی اگر $Q.W_1 \leq Q.W_2$ آنگاه رابطه $C_Q.C_{W_1} \leq C_Q.C_{W_2}$ برقرار است، به‌طوری که C_{W_1}، C_{W_2} و C_Q به ترتیب بردارهای رمز شده مربوط به بردارهای اصلی W_1، W_2 و Q هستند و عملگر "." نشان‌دهنده ضرب داخلی است. در این مفهوم رمزگذاری که از فرایندهایی همچون تقسیم تصادفی، ضرب ماتریس و افزودن بعد ساختگی استفاده می‌شود، بردارهای اصلی را نمی‌توان بدون داشتن کلید خصوصی و با تجزیه‌وتحلیل متن رمز شده متناظر آن‌ها بازیابی نمود. از طرف دیگر، ASPE در برابر حمله با متن اصلی منتخب مقاوم است، به همین جهت از آن برای محاسبه حاصل‌ضرب داخلی دو بردار رمز شده استفاده می‌شود [۱۴] که به دلیل امن و کارا بودن، برای جست‌وجوی چند کلیدواژه در داده‌های رمز شده نیز استفاده می‌شود [۱۲, ۲۹].

۵. روش پیشنهادی

در این بخش، ابتدا روش پیشنهادی به‌تفصیل بیان شده و در ادامه، طرح رمزگذاری پیشنهادی ارائه می‌گردد.

۵-۱. مدل پیشنهادی

همان‌طور که در مثال ارائه‌شده در بخش ۳ ملاحظه شد، امتیاز پزشک شماره ۱ بیشتر از مقدار امتیاز تعیین‌شده توسط بیمار و مدت‌زمان انتظار برای پزشک شماره ۸ نیز کمتر از حد تعیین‌شده توسط بیمار است. باتوجه‌به این مهم، معرفی پزشکان ۱ و ۸ به بیمار، رضایتمندی بیشتری برای بیمار به همراه خواهد

درخواست بیمار	عمومی	تخصص				فوق تخصص										امتیاز					زمان انتظار				
		a	b	c	d	a1	a2	b1	b2	c1	c2	d1	d2	۱	۲	۳	۴	۵	۱	۲	۳	۴	۵
	۱	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱	۰	۱	۱	۰	۰	۰	۰	۰
	۱	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱	۱	۱	۱	۰	۰	۰	۰	۰

شکل (۴): تبدیل بردار درخواست بیمار به بردار جدید با توجه به مقادیر ۴ و ۲ به عنوان مقادیر آستانه امتیاز و زمان انتظار

داشت. به همین دلیل، هر ترکیبی از امتیاز بیشتر و زمان انتظار کمتر از حد تعیین‌شده توسط بیمار، می‌تواند به‌عنوان معیاری برای انتخاب پزشک مناسب در نظر گرفته شود. برای دستیابی به این هدف، در بردار درخواست بیمار در بخش امتیاز، تمامی اعداد بیشتر و در بخش زمان انتظار، تمامی اعداد کمتر از حد تعیین شده توسط بیمار برابر ۱ قرار داده می‌شوند. شکل (۴) توضیحات ارائه‌شده را به تصویر کشیده است:

باتوجه‌به بردار درخواست بیمار، تطابق از طریق سرور با دو شرط "عناوین خدمت" و "امتیاز و زمان انتظار" مشخص می‌شود. به این معنی که دو بردار مربوط به یک پزشک و یک بیمار، زمانی منطبق هستند که اولاً نوع خدمتی که پزشک ارائه می‌دهد با نوع خدمت درخواستی از سمت بیمار مطابقت داشته باشد و ثانیاً پزشک موردنظر حداقل امتیاز و حداکثر زمان انتظاری را که بیمار مشخص کرده است، داشته باشد. باید توجه کرد که در این مسئله فرض بر این است که حداقل یک پزشک منطبق با درخواست بیمار وجود دارد.

در روش پیشنهادی، سرور از جست‌وجوی شباهت بر مبنای حاصل‌ضرب داخلی بردار برای یافتن مجموعه‌ای از پزشکانی استفاده می‌کند که بیشترین شباهت به بردار درخواست بیمار را دارند. خاطرنشان می‌شود باتوجه‌به اهمیت و لزوم حفظ حریم خصوصی کاربران و امنیت داده‌ها، می‌بایست قبل از بارگذاری داده‌ها روی سرور آن‌ها را رمزگذاری کرده و اطلاعات به‌صورت رمز شده روی سرور بارگذاری شوند. درصورتی‌که هر یک از پزشکان، اطلاعات مربوط به خود را ازروی‌عمد، به‌اشتباه رمزگذاری و بارگذاری نماید، به‌محض گزارش‌دهی از سمت بیماران، پس از صحت‌سنجی ادعای مطرح شده، پزشک موردنظر از سیستم حذف می‌گردد.

به‌منظور حفظ حریم خصوصی کاربران، قبل از بارگذاری اطلاعات روی سرور، می‌بایست بردارهای دودویی "نوع خدمت" و "امتیاز و زمان انتظار" تمام پزشکان، نمایه و درخواست بیمار به‌صورت رمز شده تولید شود. روش ارائه شده، باتوجه‌به درخواست رمز شده بیمار، مجموعه‌ای از مناسب‌ترین پزشکان را بدون رمزگشایی و با استفاده از نمایه و دریچه رمز شده، برمی‌گرداند. جزئیات روش ارائه‌شده در ادامه آورده شده است.

۵-۲. طرح رمزگذاری پیشنهادی

پیش از ارائه جزئیات روش، لازم به ذکر است که باتوجه‌به ماهیت ضرب داخلی و به‌منظور کسب نتایج قابل‌استناد، نیاز به بررسی بردارهای W و V به‌صورت مجزا است. بااین‌وجود، روند رمزگذاری و بررسی تطابق برای هر دو بردار W و V مشابه یکدیگر است، لذا، به‌منظور جلوگیری از بروز ابهام و افزایش پیچیدگی متن، روش ارائه شده در تمامی الگوریتم‌ها تنها برای بردار W بیان شده است که به طور مشابه برای بردار V نیز اجرا می‌شود. علاوه بر این، باتوجه‌به اینکه طول بردار W و V با یکدیگر متفاوت است، مقدار l که نشان‌دهنده طول بردار است در هر یک از الگوریتم‌ها متناسب با بردار ورودی، تعیین می‌شود؛ یعنی در الگوریتم‌های ارائه شده مقدار l برای بردار W نشان‌دهنده تعداد کل تخصص‌ها، فوق‌تخصص‌ها و مؤلفه عمومی است و برای بردار V نشان‌دهنده اندازه بردار امتیاز و زمان انتظار یعنی ۱۰ است.

نمادهای استفاده شده در این بخش در جدول (۴) آورده شده است.

جدول (۴): نمادهای استفاده شده به همراه توضیحات

نماد	توضیحات
	پارامتر امنیتی
l	طول بردار ورودی
M_1, M_2	ماتریس معکوس‌پذیر $(3l+4)\times(3l+4)$ بعدی
R	اعداد حقیقی
S	بردار $(3l+4)$ بعدی
π	تابع جای‌گشت
msk	کلید خصوصی اصلی
sk	کلید خصوصی
rk	کلید باز رمزگذاری
$M_{d,1}, M_{d,2}$	ماتریس $(3l+4)\times(3l+4)$بعدی معکوس پذیر برای تولید کلید خصوصی پزشکان
$M'_{d,1}, M'_{d,2}$	ماتریس $(3l+4)\times(3l+4)$بعدی معکوس‌پذیر برای تولید کلید باز رمزگذاری پزشکان
$M_{q,2}, M_{q,1}$	ماتریس $(3l+4)\times(3l+4)$بعدی معکوس پذیر برای تولید کلید خصوصی بیماران
$M'_{q,1}, M'_{q,2}$	ماتریس $(3l+4)\times(3l+4)$بعدی معکوس‌پذیر برای تولید کلید باز رمزگذاری بیماران
W'	بردار گسترش‌یافته W
$\widehat{W}$	بردار گسترش‌یافته W'
α, r_b, β, r_q	اعداد تصادفی مثبت یک‌بارمصرف[1]
$\lvert W'\rvert^2$	$W'.W'$
$\widetilde{W}$	$\pi(\widehat{W})$
$\tilde{C}_W$	بردار رمز شده $\widetilde{W}$ (نمایه)
C_W	بردار باز رمزگذاری شده $\tilde{C}_W$ (تبدیل نمایه)
R_{temp}	نتایج اولیه تطبیق
R	نتایج نهایی تطبیق

[1] One-Time Random Positive Numbers

- **راه‌اندازی l $(1^\lambda$,**:

مرکز تولید کلید (KGC) با استفاده از پارامتر امنیتی λ دو ماتریس $(3l+4)\times(3l+4)$ بعدی معکوس‌پذیر M_1 و M_2، بردار $(3l+4)$ بعدی S و تابع جای‌گشت $\pi: R^{3l+4} \to R^{3l+4}$ را به‌صورت تصادفی تولید می‌کند. در نهایت کلید خصوصی اصلی به‌صورت زیر تعریف می‌شود:

$$msk = \{M_1, M_2, S, \pi\} \quad (۱)$$

شبه‌کد (۱): راه‌اندازی

ورودی: پارامتر امنیتی λ

خروجی: کلید خصوصی اصلی (msk)

۱. دو ماتریس معکوس‌پذیر تصادفی M_1 و M_2با ابعاد $(3l+4)\times(3l+4)$ تولید کن.

۲. یک بردار تصادفی S با طول $(3l+4)$ تولید کن.

۳. یک تابع جای‌گشت تصادفی π که $R\hat{}(3l+4)$ را به $R\hat{}(3l+4)$ نگاشت می‌کند، تولید کن.

۴. کلید خصوصی اصلی را به‌صورت زیر تعریف کن:

$$msk = \{M_1, M_2, S, \pi\}$$

۵. مقدار msk را برگردان.

- **تولید کلید $(\boldsymbol{msk}, \boldsymbol{d_i}/\boldsymbol{q_j})$**:

این الگوریتم، کلید خصوصی اصلی و شناسه کاربر را به‌عنوان ورودی دریافت می‌کند. کاربر می‌تواند پزشک و یا بیمار باشد که کلید خصوصی و کلید باز رمزگذاری هر یک به‌صورت زیر تولید می‌شود:

پزشک i–ام $(\boldsymbol{d_i})$: ابتدا KGC به‌صورت تصادفی دو ماتریس $(3l+4)\times(3l+4)$ بعدی معکوس‌پذیر $M_{d_i,2}$و $M_{d_i,1}$ را تولید کرده و سپس ماتریس‌های زیر را محاسبه می‌کند:

$$M'_{di,1} = M_{di,1}^{-1} M_1$$
$$M'_{di,2} = M_{di,2}^{-1} M_2 \quad (۲)$$

در ادامه KGC کلیدهای sk_i و rk_i را برابر با مقادیر زیر قرار می‌دهد:

$$sk_i = \{M_{di,1}, M_{di,1}, \mathrm{S}, \pi\}$$
$$rk_i = \{M'_{di,1}, M'_{di,1}\} \quad (۳)$$

بیمار j–ام $(\boldsymbol{q_j})$: KGC ابتدا به‌صورت تصادفی دو ماتریس $(3l+4)\times(3l+4)$ بعدی معکوس‌پذیر $M_{q_j,1}$ و $M_{q_j,2}$ را تولید کرده و سپس ماتریس‌های زیر را محاسبه می‌کند:

$$M'_{q_j,1} = M_1^{-1} M^{-1}{}_{q_j,1} \quad (۴)$$

$$M'_{q_j,2} = M_2^{-1}M^{-1}{}_{q_j,2}$$

در ادامه KGC کلیدهای sk_j و rk_j را برابر با مقادیر زیر قرار می‌دهد:

$$sk_j = \{M_{q_j,1}, M_{q_j,2}, \mathrm{S}, \pi\}$$
$$rk_j = \{M'_{q_j,1}, M'_{q_j,2}\} \quad (۵)$$

در نهایت، کلیدهای خصوصی تولید شده به کاربران و کلیدهای باز رمزگذاری به سرور ارسال می‌شوند.

شبه‌کد (۲): تولید کلید

ورودی: کلید خصوصی اصلی (msk)، شناسه کاربر (id)
خروجی: کلید خصوصی (sk) و کلید باز رمزگذاری (rk)
۱. اگر کاربر i-ام پزشک باشد (d_i):
۱،۱. دو ماتریس معکوس‌پذیر تصادفی $M_{d_i,1}, M_{d_i,2}$ را به ابعاد $(3l+4)\times(3l+4)$ تولید کن.
۱،۲. ماتریس‌های $M'_{d,2}$ و $M'_{d,1}$ را به‌صورت زیر محاسبه کن: $M'_{di,1} = M_{di,1}^{-1}M_1$ ، $M'_{di,2} = M_{di,2}^{-1}M_2$
۱،۳. کلید خصوصی و باز رمزگذاری را به‌صورت زیر تعیین کن: $sk_i = \{M_{di,1}, M_{di,1}, \mathrm{S}, \pi\}$ ، $rk_i = \{M'_{di,1}, M'_{di,1}\}$
۲. اگر کاربر j-ام بیمار باشد (q_j):
۲،۱. دو ماتریس معکوس‌پذیر تصادفی $M_{q,2}, M_{q,1}$ را با ابعاد $(3l+4)\times(3l+4)$ تولید کن.
۲،۲. ماتریس‌های زیر را محاسبه کن: $M'_{q_j,1} = M_1^{-1}M^{-1}{}_{q_j,1}$ ، $M'_{q_j,2} = M_2^{-1}M^{-1}{}_{q_j,2}$
۲،۳. کلید خصوصی و باز رمزگذاری را به‌صورت زیر تعیین کن: $sk_j = \{M_{q_j,1}, M_{q_j,2}, \mathrm{S}, \pi\}$ ، $rk_j = M'_{q_j,1}, M'_{q_j,2}$
۳. کلید خصوصی به کاربر و کلید باز رمزگذاری را به سرور ارسال کن.

- **رمزگذاری نمایه: (sk_i, W_i)**

به‌منظور ساخت نمایه، عملیات زیر روی بردار پزشکان انجام می‌شود:

گام ۱: بردار ورودی به طول l با استفاده از عملیات زیر به بردار به طول $2l$ گسترش می‌یابد:

$$W'_{i,2j-1} = W_{i,j}\ ,\ \ j\epsilon[1,l]$$
$$W'_{i,2j} = \begin{cases} 0\ , & j\epsilon[1,\frac{l}{2}] \\ 1, & j\epsilon(\frac{l}{2},l] \end{cases} \quad (۶)$$

گام ۲: بردار $3l+4$ بعدی $\widehat{W_i} = (\alpha W_i', \alpha W, \alpha, \alpha|W_i'|^2, r_i, 0)$ به‌طوری که α و r_i اعداد تصادفی مثبت یک‌بارمصرف هستند و $|W_i'|^2 = W_i'.W_i'$ ساخته می‌شوند.

گام ۳: بردار $\widetilde{W_i}$ با استفاده از $\widetilde{W_i} = \pi(\widehat{W_i})$ به دست می‌آید.

گام ۴: بردار $\widetilde{W_i}$ به دو بردار تصادفی $\{\widetilde{W_i}^a, \widetilde{W_i}^b\}$ تقسیم شده و برای z از ۱ تا $3l+4$، هر مؤلفه به‌صورت زیر تنظیم می‌شود:

$$\begin{cases} \widetilde{W_i}^a[z] = \widetilde{W_i}^b[z] = \widetilde{W_i}[z], & S[z]=0 \\ \widetilde{W_i}^a[z] + \widetilde{W_i}^b[z] = \widetilde{W_i}[z], & S[z]=1 \end{cases} \quad (۷)$$

گام ۵: بردارهای $\{\widetilde{W_i}^a, \widetilde{W_i}^b\}$ به‌صورت $\{M_{i,1}^T\widetilde{W_i}^a, M_{i,2}^T\widetilde{W_i}^b\}$ رمزگذاری شده و به همراه شناسه پزشک، روی سرور بارگذاری می‌گردد.

شبه‌کد (۳): رمزگذاری نمایه

ورودی: کلید خصوصی (sk)، بردار کاربر (W)
خروجی: بردار رمز شده ($\tilde{C}_W$)
۱. بردار ورودی W_i را به دوبرابر طول اولیه‌اش به‌صورت زیر گسترش بده: $W'_{i,2j-1} = W_{i,j}\ ,\ \ j\epsilon[1,l]$ ، $W'_{i,2j} = \begin{cases} 0\ , & j\epsilon[1,\frac{l}{2}] \\ 1, & j\epsilon(\frac{l}{2},l] \end{cases}$
۲. بردار گسترش‌یافته $\widehat{W_i}$ را با استفاده از اعداد تصادفی یک‌بارمصرف α و r_b به‌صورت زیر بساز: $\widehat{W_i} = (\alpha W_i', \alpha W, \alpha, \alpha\|W_i'\|^2, r_i, 0)$
۳. جای‌گشت تصادفی زیر را روی بردار به‌دست‌آمده از گام قبل اعمال کن: $\widetilde{W_i} = \pi(\widehat{W_i})$
۴. بردار $\widetilde{W_i}$ را به دو بخش تصادفی برای z از ۱ تا $3l+4$ به‌صورت زیر تقسیم کن: $\widetilde{W_i}^a, \widetilde{W_i}^b = \begin{cases} \widetilde{W_i}^a[z] = \widetilde{W_i}^b[z] = \widetilde{W_i}[z], & S[z]=0 \\ \widetilde{W_i}^a[z] + \widetilde{W_i}^b[z] = \widetilde{W_i}[z], & S[z]=1 \end{cases}$
۵. نمایه $\tilde{C}_{W_i}$ را به‌صورت زیر رمزگذاری کن: $\tilde{C}_{W_i} = \{M_{i,1}^T\widetilde{W_i}^a, M_{i,2}^T\widetilde{W_i}^b\}$
۶. $\tilde{C}_{W_i}$ را روی سرور ذخیره کن.

- **تبدیل نمایه**[1] $(rk_i, \tilde{C}_{W_i})$:

زمانی که سرور، متن رمز شده $\tilde{C}_{W_i}$ را برای هر پزشک دریافت می‌کند، ابتدا کلید rk_i مربوطه را پیدا کرده و متن رمز شده را به‌صورت زیر باز رمزگذاری می‌کند:

$$C_{W_i} = \{M'^T_{i,1} M^T_{i,1} \widetilde{W}^a_i, M'^T_{i,2} M^T_{i,2} \widetilde{W}^b_i\} = \{M^T_1 \widetilde{W}^a_i, M^T_2 \widetilde{W}^b_i\} \quad (۸)$$

شبه‌کد (۴): تبدیل نمایه

ورودی: کلید باز رمزگذاری (rk_i)، بردار رمز شده نمایه $(\tilde{C}_{W_i})$ خروجی: بردار باز رمزگذاری شده (C_{W_i})
۱. کلید باز رمزگذاری rk_i را دریافت کن. ۲. تبدیل زیر را روی بردار $\tilde{C}_{W_i}$ اعمال کن: $C_{W_i} = \{M'^T_{i,1} M^T_{i,1} \widetilde{W}^a_i, M'^T_{i,2} M^T_{i,2} \widetilde{W}^b_i\} = \{M^T_1 \widetilde{W}^a_i, M^T_2 \widetilde{W}^b_i\}$ ۳. بردار باز رمزگذاری شده C_{W_i} را روی سرور ذخیره و نگهداری کن.

- **تولید دریچه** (sk_j, W_j):

با دریافت کلید خصوصی sk_j و بردار درخواست بیمار q_j، بردار رمز شده مربوط به بیمار طبق مراحل زیر تولید می‌شود:

گام ۱: ابتدا بردار l-بعدی W_j به بردار 2l بعدی $W'_j = (W'_{j,1}, W'_{j,2}, \dots, W'_{j,2l-1}, W'_{j,2l})$ به‌صورت زیر گسترش می‌یابد:

$$\begin{cases} W'_{j,2i-1} = W_j, & i\epsilon[1,l] \\ W'_{j,2i} = 1, & i\epsilon[1,l] \end{cases} \quad (۹)$$

گام ۲: W'_j با استفاده از اعداد تصادفی مثبت یک‌بارمصرف β, r_j به بردار $3l+4$ بعدی زیر تبدیل می‌شود:

$$\widehat{W_j} = ((1+r_j)\beta W'_j, -r_j\beta W_j, -r_j\beta|W'_j|^2, -r_j\beta, 0, r_j)$$

به‌طوری که $|W'_j|^2 = W'_j . W'_j$.

گام ۳: بردار $\widehat{W_j}$ با استفاده از $\widehat{W_j} = \pi(\widehat{W_j})$ محاسبه می‌شود.

گام ۴: بردار $\widehat{W_j}$ به دو بردار $\{\widetilde{W}^a_j, \widetilde{W}^b_j\}$ تقسیم می‌شود، به‌طوری که برای z از ۱ تا $3l+4$، هر بیت به‌صورت زیر تنظیم می‌شود:

$$\begin{cases} \widetilde{W}^a_j[z] + \widetilde{W}^b_j[z] = \widehat{W_j}[z], & S[z] = 0 \\ \widetilde{W}^a_j[z] = \widetilde{W}^b_j[z] = \widehat{W_j}[z], & S[z] = 1 \end{cases} \quad (۱۰)$$

گام ۵: در نهایت با استفاده از کلید خصوصی sk_j، $\{\widetilde{W}^a_j, \widetilde{W}^b_j\}$ به‌صورت $\{M^{-1}{}_{j,1}\widetilde{W}^a_j, M^{-1}{}_{j,2}\widetilde{W}^b_j\}$ رمز می‌شود. در این حالت، بردار W_j رمز شده درخواست بیمار به شکل $\widetilde{T_{W_j}} = \{M^{-1}_{j,1}\widetilde{W}^a_j, M^{-1}_{j,2}\widetilde{W}^b_j\}$ نشان داد می‌شود. در همین راستا بیمار $\widetilde{T_{W_j}}$ مربوطه را به سرور ارسال می‌نماید.

شبه‌کد (۵): تولید دریچه

ورودی: کلید خصوصی بیمار (sk_j)، بردار درخواست بیمار (q_j) خروجی: دریچه رمزگذاری شده $(\widetilde{T_{W_j}})$
۱. بردار ورودی W_j را به دوبرابر طول اولیه به‌صورت زیر گسترش بده: $W'_j = \begin{cases} W'_{j,2i-1} = W_j, & i\epsilon[1,l] \\ W'_{j,2i} = 1, & i\epsilon[1,l] \end{cases}$ ۲. بردار گسترش‌یافته $\widehat{W_j}$ را با استفاده از اعداد تصادفی یک‌بارمصرف β و r_j به‌صورت زیر بساز: $\widehat{W_j} = ((1+r_j)\beta W'_j, -r_j\beta W_j, -r_j\beta|W'_j|^2, -r_j\beta, 0, r_j)$ ۳. جای‌گشت تصادفی زیر را روی بردار به‌دست‌آمده از گام قبل اعمال کن: $\widehat{W_j} = \pi(\widehat{W_j})$ ۴. بردار $\widehat{W_j}$ را به دو بخش تصادفی برای z از ۱ تا $3l+4$ به‌صورت زیر تقسیم کن: $\widetilde{W}^a_j, \widetilde{W}^b_j = \begin{cases} \widetilde{W}^a_j[z] + \widetilde{W}^b_j[z] = \widehat{W_j}[z], & S[z] = 0 \\ \widetilde{W}^a_j[z] = \widetilde{W}^b_j[z] = \widehat{W_j}[z], & S[z] = 1 \end{cases}$ ۵. بردار $\widetilde{T_{W_j}}$ را به‌صورت زیر رمزگذاری کن: $\widetilde{T_{W_j}} = \{M^{-1}_{j,1}\widetilde{W}^a_j, M^{-1}_{j,2}\widetilde{W}^b_j\}$ ۶. $\widetilde{T_{W_j}}$ را روی سرور ذخیره کن.

- **تبدیل دریچه** $(rk_j, \widetilde{T_{W_j}})$:

زمانی که سرور، متن رمز شده $\widetilde{T_{W_j}}$ را دریافت می‌کند، ابتدا کلید rk_j مربوطه را پیدا کرده و T_{W_j} را به‌صورت زیر محاسبه می‌کند:

[1] Index Transform

$$T_{W_j} = \{M'_{j,1}M^{-1}{}_{j,1}\widetilde{W}_j^a, M'_{j,2}M_{j,2}^{-1}\widetilde{W}_j^b\} = \{M_1^{-1}\widetilde{W}_j^a, M_2^{-1}\widetilde{W}_j^b\} \quad (۱۱)$$

شبه‌کد (۶): تبدیل دریچه

ورودی: کلید باز رمزگذاری بیمار (rk_j)، متن رمز شده دریچه ($\widetilde{T_{W_j}}$)
خروجی: دریچه تبدیل‌شده (T_{W_j})
۱. گام ۱: کلید باز رمزگذاری rk_j را دریافت کن. ۲. گام ۲: تبدیل زیر را روی بردار $\widetilde{T_{W_j}}$ اعمال کن: $T_{W_j} = \{M'_{j,1}M^{-1}{}_{j,1}\widetilde{W}_j^a, M'_{j,2}M_{j,2}^{-1}\widetilde{W}_j^b\} = \{M_1^{-1}\widetilde{W}_j^a, M_2^{-1}\widetilde{W}_j^b\}$ ۳. بردار باز رمزگذاری شده T_{W_j} را روی سرور ذخیره و نگهداری کن.

- **جست‌وجو ($C_{W_i}, C_{V_i}, T_{W_j}, T_{V_j}$):**

در این مرحله، سرور عملیات تطبیق را برای T_{W_j} و C_{W_i} انجام می‌دهد و با استفاده از فرایند زیر بررسی می‌کند که آیا بردارهای W_i و W_j با یکدیگر تطابق دارند یا خیر:

گام ۱: در ابتدا، ضرب داخلی T_{W_j} و C_{W_i} به‌صورت زیر توسط سرور محاسبه می‌شود:

$$\begin{aligned} C_{W_i}.T_{W_j} &= \{M_1^T\widetilde{W}_j^a, M_2^T\widetilde{W}_j^b\} \circ \{M_1^{-1}\widetilde{W}_j^a, M_2^{-1}\widetilde{W}_j^b\} \\ &= \{M_1^T\widetilde{W}_j^a, M_2^T\widetilde{W}_j^b\}^{\boldsymbol{T}}\{M_1^{-1}\widetilde{W}_j^a, M_2^{-1}\widetilde{W}_j^b\} \\ &= \widetilde{W_i^a}^T\widetilde{W}_j^a + \widetilde{W_i^b}^T\widetilde{W}_j^b \\ &= \widetilde{W_i}.\widetilde{W_j} \end{aligned} \quad (۱۲)$$

پس از محاسبه مقادیر $C_{W_i}.T_{W_j}$، نتایج با بیشترین مقدار نشان می‌دهند که دو بردار مربوطه بیشترین شباهت را به یکدیگر دارند. سرور شناسه مربوط به پزشکانی را که بیشترین مقادیر را دارند به‌عنوان نتایج اولیه به مجموعه نتایج R_{temp} اضافه می‌کند:

$$R_{temp} = \{i|MAX(C_{W_i}.T_{W_j})\} \quad (۱۳)$$

پس از به‌دست‌آوردن شناسه پزشکانی که بردار "نوع خدمت" آن‌ها با بخش مربوطه در درخواست بیمار منطبق است و کنارگذاشتن باقی پزشکان، به بررسی بردار "امتیاز و زمان انتظار" پرداخته می‌شود. به این منظور، مشابه قبل، پس از محاسبه $C_{V_i}.T_{V_j}$ برای نتایج موجود در R_{temp}، نتایج با بیشترین مقدار به این معنی است که بردار "امتیاز و زمان انتظار" پزشک منتخب، مشابه‌ترین بردار به بردار درخواست بیمار است و سرور، شناسه پزشک موردنظر را به مجموعه نهایی R جهت برگرداندن به بیمار اضافه می‌کند:

$$R = \{i|MAX(C_{V_i}.T_{V_j})\} \quad (۱۴)$$

شبه‌کد (۷): جست‌وجو

ورودی: بردار رمزگذاری شده پزشکان (C_{W_i} و C_{V_i})، دریچه رمزگذاری شده بیمار (T_{W_j} و T_{V_j})
خروجی: پزشکان با بیشترین تطابق (مجموعه R)
۱. شباهت بین بردارهای C_{W_i} و T_{W_j} برای هر پزشک i موجود در پایگاه‌داده را با استفاده از ضرب داخلی به‌صورت زیر محاسبه کن: $C_{W_i}.T_{W_j} = \{M_1^T\widetilde{W}_j^a, M_2^T\widetilde{W}_j^b\} \circ \{M_1^{-1}\widetilde{W}_j^a, M_2^{-1}\widetilde{W}_j^b\}$ $= \{M_1^T\widetilde{W}_j^a, M_2^T\widetilde{W}_j^b\}^{\boldsymbol{T}}\{M_1^{-1}\widetilde{W}_j^a, M_2^{-1}\widetilde{W}_j^b\}$ $= \widetilde{W_i^a}^T\widetilde{W}_j^a + \widetilde{W_i^b}^T = \widetilde{W_i}.\widetilde{W_j}$ ۲. پزشکان با بیشترین مقدار شباهت در بخش W را انتخاب کن و نتایج را در بردار موقت R_{temp} ذخیره کن: $R_{temp} = \{i|MAX(C_{W_i}.T_{W_j})\}$ ۳. شباهت بین بردارهای C_{V_i} و T_{V_j} را برای هر پزشک i موجود در R_{temp} با استفاده از ضرب داخلی مشابه گام قبل محاسبه کن: $C_{V_i}.T_{V_j} = \widetilde{V_i}.\widetilde{V_j}$ ۴. پزشکان با بیشترین مقدار شباهت در بخش W را انتخاب کن و نتایج را در بردار R ذخیره کن: $R = \{i|MAX(C_{V_i}.T_{V_j})\}$ ۵. مقدار R را به‌عنوان خروجی برگردان.

۶. تحلیل درستی روش پیشنهادی

در این بخش درستی روش ارائه شده مورد بررسی قرار می‌گیرد. طبق تعاریف ارائه‌شده در بخش‌های قبل، قضیه زیر برقرار است:

قضیه ۱: نتیجه برگردانده شده توسط روش ارائه‌شده صحیح است، یعنی مجموعه‌ها ی $R_{temp} = \{i|MAX(C_{W_i}.T_{W_j})\}$ و $R = \{i|MAX(C_{V_i}.T_{V_j})\}$ شامل پزشکانی هستند که بیشترین مقدار نتیجه ضرب داخلی با بردار درخواست بیمار را داشته‌اند. به بیان دیگر هیچ پزشکی وجود ندارد که مقدار نتیجه ضرب داخلی بردار آن با بردار درخواست بیمار بیشتر از مقدار نتیجه ضرب داخلی پزشکانی باشد که به بیمار برگردانده می‌شوند و در مجموعه R_{temp} یا R قرار گرفته‌اند.

اثبات ۱-۱: درستی روش ارائه شده

باتوجه‌به نتایج برگردانده شده توسط مدل پیشنهادی، مجموعه‌های R_{temp}وR به ترتیب دارای شناسه مربوط به پزشکانی است که بیشترین مقادیر ضرب داخلی با بخش "نوع خدمت" بردار درخواست بیمار و بخش "امتیاز و زمان انتظار" بردار درخواست بیمار را دارند.

مفهوم درست بودن نتایج برگردانده شده توسط مدل پیشنهادی به این معناست که تمام پزشکانی که بیشترین شباهت به بردار درخواست بیمار را دارند، به‌عنوان خروجی برگردانده می‌شوند. به بیان دیگر، پزشک دیگری که نتیجه ضرب داخلی بردار آن با بردار درخواست بیمار، بیشتر از پزشکان برگردانده شده باشد، وجود ندارد که عضو مجموعه‌های R_{temp}و R نباشند.

با استفاده از برهان خلف، قضیه فوق قابل‌اثبات است. فرض می‌شود که حداقل یک پزشک وجود دارد که مقدار ضرب داخلی بردار آن با بردار درخواست بیمار بیشتر از مقدار ضرب داخلی بردارهایی است که به‌عنوان خروجی در مجموعه‌های R_{temp}یا R برگردانده شده است، لیکن به‌عنوان پزشک مناسب در خروجی وجود ندارد. در این صورت، بردار مذکور می‌بایست در گام دوم الگوریتم جست‌وجو که پزشکان با بیشترین مقدار ضرب داخلی انتخاب شده و در مجموعه‌های R_{temp} یا R قرار می‌گرفت. حال که این انتخاب صورت نگرفته است، در نتیجه مجموعه‌های R_{temp} یا R بیشینه مقادیر حاصل‌ضرب داخلی بردارها را شامل نبوده و این خلاف فرض مسئله است، لذا فرض خلف باطل و هیچ بردار پزشکی وجود ندارد که مقدار ضرب داخلی آن با بردار درخواست بیمار بیشتر از مقدار برداری باشد که به‌عنوان خروجی برگردانده شده است و درستی روش پیشنهادی اثبات می‌گردد.

۷. تجزیه‌وتحلیل امنیتی

در تجزیه‌وتحلیل امنیتی، امنیت روش ارائه شده دررابطه‌با مدل تهدید و الزامات امنیتی بیان شده در بخش ۳، مورد بررسی قرار می‌گیرد. لازم به ذکر است به علت تشابه روش رمزگذاری هر دو بردار "نوع خدمت" و "امتیاز و زمان انتظار" تحلیل‌های ارائه شده تنها برای یک بخش از بردارها صورت پذیرفته و قابل‌تعمیم به بخش دیگر است. همچنین همان‌طور که در مدل تهدید توضیح داده شد، در مدل CPA، سرور اطلاعات بیشتری نسبت به مدل KPA دارد. یعنی اگر روش ارائه شده تحت مدل CPA ایمن باشد، تحت مدل KPA نیز امن است؛ بنابراین، در ادامه از اثبات امنیت تحت مدل KPA صرف‌نظر شده و روی اثبات امنیت تحت مدل CPA تمرکز می‌شود.

قضیه ۲: روش ارائه شده محرمانگی نمایه‌ها و محرمانگی دریچه-ها را تحت مدل CPA فراهم می‌کند.

اثبات ۲-۱: محرمانگی نمایه

نمادهای i، W_i و $\tilde{C}_{W_i}$ به ترتیب نشان‌دهنده شناسه پزشک، بردار مرتبط با پزشک i-ام و بردار رمز شده W_i است. در مدل CPA، سرور می‌تواند تعدادی $\tilde{C}_{W_i}$ و بردار دودویی مربوطه W_i را به انتخاب خود انتخاب نماید. بدون داشتن بردار S و تابع π، سرور نمی‌تواند کلید خصوصی $sk_i = \{M_{i,1}, M_{i,2}\}$ را بازیابی کند.

بردار W_i در الگوریتم رمزگذاری نمایه به یک بردار $3l+4$ بعدی $\widetilde{W_i} = (\alpha W_i', \alpha W, \alpha, \alpha|W_i'|^2, r_i, 0)$ تبدیل می‌شود. امنیت اعداد مثبت یک‌بارمصرف α و r_i که به‌صورت تصادفی هستند، برابر با امنیت رمزگذاری پد یک‌بارمصرف [1] است که می‌تواند امنیت کامل [2] را فراهم آورد و این اطمینان را بدهد که C_{W_i} یک بردار تصادفی برای سرور است.

همچنین باتوجه‌به اینکه هر دو ماتریس M_1 و M_2 برای سرور ناشناخته است و کلید خصوصی $sk_i = \{M_{i,1}, M_{i,2}\}$ برای هر یک از پزشکان متفاوت است، سرور نمی‌تواند راه‌حل دقیقی برای معادلات خطی زیر به دست آورد:

$$\begin{matrix} \widetilde{C_{W_i}^a} \\ = M_1^T \widetilde{W}_i^a \\ \widetilde{C_{W_i}^b} \\ = M_2^T \widetilde{W}_i^b \end{matrix} \quad (۱۵)$$

در دستگاه معادلات فوق، تعداد مجهولات در بردارهای $\widetilde{W}_j^a$و $\widetilde{W}_j^b$ و ماتریس‌های M_1 و M_2 به ترتیب برابر $2(3l+4)$ و $2(3l+4)^2$ است. بااین‌حال، تنها $2(3l+4)$ معادله وجود دارد که تعداد آن‌ها بسیار کمتر از تعداد مجهولات است؛ بنابراین، سرور نمی‌تواند یک حمله تحلیل خطی [3] را با اطلاعات به‌دست‌آمده در مدل CPA اجرا نماید.

علاوه‌برآن، سرور تنها می‌داند که حاصل‌ضرب داخلی $C_{W_i}.C_{W_j}$ چه رابطه‌ای با حاصل‌ضرب داخلی $C_{W_i}.C_{W_j}$ دارد (بزرگ‌تر، مساوی یا کوچک‌تر) به‌طوری که $i \neq i$ که این امر نشان می‌دهد تعداد کلیدواژه‌های یکسان بین بردار پزشک و بردار درخواست بیمار چند تا است و هیچ ایده‌ای در مورد اطلاعات کلیدی در خصوص W_j،W_i یا $W_j.W_i$ندارد.

اثبات ۲-۲: محرمانگی دریچه

ازآنجاکه ساختار تولید دریچه مشابه ساختار تولید نمایه است، مطابق اثبات محرمانگی نمایه، به‌راحتی می‌توان اثبات نمود که محرمانگی دریچه‌ها تحت مدل CPA فراهم است.

قضیه ۳: روش ارائه شده می‌تواند عدم پیوند نمایه و عدم پیوند دریچه را تحت مدل CPA فراهم کند.

[1] One-Time Pad
[2] Perfect Security
[3] Linear Analysis Attack

اثبات ۳-۱: عدم پیوند نمایه

فرض کنید که W_i و $W_{i'}$ دو نمایش برداری یکسان برای دو پزشک متفاوت است. تصادفی‌سازی‌های مورد استفاده در الگوریتم رمزگذاری نمایه را می‌توان در سه حوزه‌ی پد یک‌بارمصرف تصادفی [1]، جای‌گشت تصادفی [2] و تقسیم‌بندی تصادفی [3] مورد بررسی قرارداد.

- پد یک‌بارمصرف تصادفی: همان‌طور که در الگوریتم رمزگذاری نمایه گفته شد، دو بردار W_i و $W_{i'}$ به شکل زیر گسترش می‌یابند:

$$\begin{matrix}\widetilde{W_i} \\ = (\alpha W_i', \alpha W, \alpha, \alpha|W_i'|^2, r_i, 0) \\ \widetilde{W_{i'}} \\ = (\alpha W_{i'}', \alpha W, \alpha, \alpha|W_{i'}'|^2, r_i, 0)\end{matrix} \quad (۱۶)$$

ازآنجاکه α، α'، r_i و r'_i اعداد تصادفی یک‌بارمصرف متفاوتی هستند، تابع رمزگذاری نمایه برای سرور، تصادفی است. به این معنی که حتی با داشتن بردارهای W_i و $W_{i'}$ به انتخاب خود سرور، متن رمز شده آن‌ها تصادفی و با یکدیگر متفاوت خواهد بود.

- جای‌گشت تصادفی: همان‌طور که در الگوریتم رمزگذاری نمایه نشان داده شد، جای‌گشت دو بردا W_i و $W_{i'}$ به‌صورت زیر محاسبه می‌شود:

$$\begin{matrix}\widetilde{W_i} \\ = \pi(\widetilde{W_i}) \\ \widetilde{W_{i'}} \\ = \pi(\widetilde{W_{i'}})\end{matrix} \quad (۱۷)$$

تابع جای‌گشت تصادفی π می‌تواند یک ترکیب جدید از عناصر دو بردار تولید کند، به‌طوری که ترتیب اصلی بردارهای توسعه‌یافته به‌صورت تصادفی تغییر یابند. با این تفاسیر سرور نمی‌تواند استنباط کند که W_i و W_i یکسان است یا حاوی کلمات کلیدی یکسانی است.

- تقسیم‌بندی تصادفی: همان‌طور که گفته شد، زمانی که مقدار $S[z]$ برابر یک است، هر دو بردار $\widetilde{W}_i^a[z]$ و $\widetilde{W}_i^b[z]$ به طور تصادفی به دو مقدار تقسیم می‌شوند که مجموع این دو مقدار تصادفی برابر مقدار اصلی است، یعنی:

$$\begin{matrix}\widetilde{W}_i^a[z] + \widetilde{W}_i^b[z] = \widetilde{W}_i[z] \\ \widetilde{W}_{i'}^a[z] + \widetilde{W}_{i'}^b[z] = \widetilde{W}_{i'}[z]\end{matrix} \quad (۱۸)$$

فرض کنید:

$$\begin{matrix}Pr[\{\widetilde{W}_i^a[z], \widetilde{W}_i^b[z]\} \\ = \{\widetilde{W}_{i'}^a[z], \widetilde{W}_{i'}^b[z]\}] \\ = \theta\end{matrix} \quad (۱۹)$$

بنابراین، برای هر بعدی از z به‌طوری که $S[z] = 1$ باشد، رابطه زیر برقرار است:

$$Pr[\{\widetilde{W}_i^a, \widetilde{W}_i^b\} = \{\widetilde{W}_{i'}^a, \widetilde{W}_{i'}^b\}] = \theta^{S_0} \quad (۲۰)$$

به‌طوری که S_0 نشان‌دهنده تعداد ۱ در بردار S است. ازآنجاکه $0 \le \theta < 1$ و $S_0 \to 3l + 4$ و هر دو C_{W_i} و $C_{W_{i'}}$ با استفاده از الگوریتم‌های رمزگذاری نمایه و تبدیل نمایه مشابه ساخته شده‌اند، احتمال زیر برقرار است:

$$Pr[C_{W_i} = C_{W_{i'}}] = Pr[W_i = W_{i'}] \cdot \theta^{S_0} \to 0 \quad (۲۱)$$

بنا بر موارد گفته شده، باتوجه‌به هر دو متن اصلی W_i و $W_{i'}$ که توسط سرور انتخاب شده‌اند، متن‌های رمز شده آن‌ها برای سرور تصادفی هستند. به این معنی که سرور نمی‌تواند تشخیص دهد که کدام بردار رمزگذاری شده است. در نتیجه ویژگی عدم پیوند نمایه تحت مدل CPA برقرار است.

اثبات ۳-۲: عدم پیوند دریچه

فرض کنید که W_j و $W_{j'}$ دو نمایش برداری یکسان برای درخواست دو بیمار متفاوت است. مشابه قسمت قبل، عدم پیوند دریچه از سه جنبه گفته شده مورد بررسی قرار می‌گیرد. ازآنجایی‌که پد یک‌بارمصرف و جای‌گشت تصادفی در الگوریتم تولید دریچه مشابه الگوریتم رمزگذاری نمایه است، از تحلیل آن صرف‌نظر شده و بر حوزه تقسیم‌بندی تصادفی در تولید دریچه تمرکز می‌شود.

باتوجه‌به دو بردار دودویی موردنیاز W_j و $W_{j'}$، همان‌طور که نشان داده شد، زمانی که $S[z]$ برابر ۰ باشد، هر دو $\widetilde{W}_j^a[z]$ و $\widetilde{W}_{j'}^a[z]$ به‌صورت تصادفی به دو مقدار تقسیم می‌شوند. یعنی

$$\begin{matrix}\widetilde{W}_j^a[z] + \widetilde{W}_j^b[z] = \widetilde{W}_j[z] \\ \widetilde{W}_{j'}^a[z] + \widetilde{W}_{j'}^a[z] = \widetilde{W}_{j'}[z]\end{matrix} \quad (۲۲)$$

فرض کنید که

$$Pr[\{\widetilde{W}_j^a[z], \widetilde{W}_j^b[z]\} = \{\widetilde{W}_{j'}^a[z], \widetilde{W}_{j'}^b[z]\}] = \eta \quad (۲۳)$$

برای تمام ابعاد z زمانی که $S[z] = 0$، رابطه زیر برقرار است:

$$Pr[\{\widetilde{W}_j^a, \widetilde{W}_j^b\} = \{\widetilde{W}_{j'}^a, \widetilde{W}_{j'}^b\}] = \eta^{S_1} \quad (۲۴)$$

به‌طوری که S_1 تعداد ۰ ها در بردار S را نشان می‌دهد.

ازآنجاکه $0 \le \eta < 1$ و $S_1 \to 3l + 4$، احتمال زیر برقرار است:

$$Pr\left[T_{W_j} = T_{W_{j'}}\right] = Pr[W_j = W_{j'}] \cdot \eta^{S_1} \to 0 \quad (۲۵)$$

[1] Random One-Time Pads
[2] Random Permutation
[3] Random Segmentation

بنابراین، باتوجه‌به هر دو بردار W_j و $W_{j'}$ که توسط سرور انتخاب شده‌اند، بردار رمز شده آن‌ها برای سرور تصادفی است. یعنی سرور نمی‌تواند تشخیص دهد که کدام دریچه رمزگذاری شده است؛ بنابراین، عدم پیوند دریچه برقرار است.

۸. ارزیابی روش پیشنهادی

در این بخش، ابتدا معیارهای ارزیابی و تنظیمات پیکربندی معرفی شده و سپس پیاده‌سازی مدل پیشنهادی و مقایسه عملکرد آن با مدل ضرب داخلی سنتی مورد بررسی قرار می‌گیرد.

۸-۱. معیارهای ارزیابی عملکرد

معیارهای ارزیابی جهت سنجش عملکرد مدل و توانایی پیش‌بینی آن مورد استفاده قرار می‌گیرند. انتخاب این معیارها می‌بایست متناسب با حوزه و هدف مسئله باشد. ازآنجا که در مسئله حاضر، از نظر بیماران مختلف معیارهای متفاوتی حائز اهمیت است، تمامی معیارهای زیر برای الگوریتم پیشنهادی و الگوریتم ضرب داخلی سنتی با یکدیگر مقایسه شده‌اند.

معیار دقت[1]: معیار دقت یکی از مهم‌ترین معیارهای ارزیابی عملکرد مدل است که نشان می‌دهد چه درصدی از پیشنهادهای ارائه‌شده توسط مدل، واقعاً مرتبط و صحیح هستند. این معیار به‌صورت زیر محاسبه می‌شود:

$$Precision = \frac{TP^{۲}}{TP + FP^{۳}} \qquad (۲۶)$$

در این مطالعه، دقت نشان‌دهنده نسبت پزشکان مناسب شناسایی‌شده به کل پزشکان پیشنهادی است. بااین‌حال، دقت به‌تنهایی ممکن است گمراه‌کننده باشد و بهتر است در کنار سایر معیارها مورد بررسی قرار گیرد.

معیار یادآوری[4]: این معیار نشان می‌دهد چه نسبتی از کل پزشکان مناسب شناسایی‌شده به‌درستی شناسایی شده‌اند. در صورت نامتوازن بودن داده‌ها، این معیار از اهمیت بیشتری برخوردار است. فرمول محاسبهٔ یادآوری به‌صورت زیر است:

$$Recall = \frac{TP}{TP + FN^{۵}} \qquad (۲۷)$$

در برخی مدل‌ها، افزایش مقدار یادآوری منجر به کاهش معیار دقت می‌شود و بالعکس؛ بنابراین، ایجاد تعادل میان این دو معیار ضروری است.

معیار امتیاز F1: این معیار به‌عنوان میانگین هارمونیک معیارهای دقت و یادآوری معرفی شده است که با استفاده از فرمول زیر میان دو معیار مذکور تعادل ایجاد می‌کند:

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall} \qquad (۲۸)$$

زمان جست‌وجو[6]: این معیار مدت‌زمان موردنیاز برای اجرای مدل را اندازه‌گیری می‌کند. در کاربردهای عملی، زمان جست‌وجوی بالا ممکن است مدل را غیرقابل‌استفاده کند، به‌ویژه هنگامی که حجم پایگاه‌داده بسیار بزرگ باشد.

۸-۲. مجموعه‌داده‌ها و پیاده‌سازی

به دلیل عدم دسترسی به پایگاه‌داده واقعی، چندین مجموعه‌داده شامل اطلاعات پزشکان با اندازه‌های مختلف و بردارهای درخواست بیماران به‌صورت تصادفی و مطابق با توضیحات بخش ۳ تولید شدند. اندازه پایگاه‌داده‌های مورد استفاده به ترتیب شامل ۱۰۰۰، ۱۵۰۰، ۲۰۰۰، ۲۵۰۰ و ۳۰۰۰ بردار پزشک است. برای بررسی معیارهای دقت، یادآوری، F1 و زمان جست‌وجو، تعداد ۱۰۰ بردار درخواست بیمار نیز به‌صورت تصادفی تولید گردید. ازآنجاکه استفاده از متغیرهای تصادفی می‌تواند نتایج متنوعی در هر اجرا ایجاد کند، هر آزمایش برای هر پایگاه‌داده پنج مرتبه اجرا شده و میانگین معیارهای ارزیابی استخراج شده‌اند. نتایج حاصل در ادامه آورده‌شده است. فرایند پیاده‌سازی با استفاده از زبان برنامه‌نویسی پایتون نسخه ۳،۸،۳ انجام شده است.

۸-۳. نتایج

در این بخش نتایج به‌دست‌آمده از پیاده‌سازی انجام شده به‌تفصیل مورد بررسی قرار گرفته است.

مقایسه نتایج معیار دقت: هرچه میزان دقت بالاتر باشد، نشان‌دهنده توانمندی سیستم در معرفی پزشکان مرتبط و مناسب به بیماران است. در مقابل، دقت پایین حاکی از آن است که تعداد زیادی از پزشکان نامرتبط به بیماران پیشنهاد شده‌اند که این امر می‌تواند منجر به اتلاف زمان بیماران و کاهش اثربخشی کلی سیستم شود.

در شرایطی که کیفیت پیشنهادها بر کمیت آن‌ها اولویت دارد به (عنوان مثال، هنگامی که بیماران نیازمند انتخاب دقیق‌تر و سریع‌تر پزشک مناسب هستند) معیار دقت نقش حیاتی‌تری ایفا می‌کند. در چنین سناریوهایی، کاهش مواجهه با گزینه‌های نامرتبط، تجربه کاربری را بهبود بخشیده و فرایند تصمیم‌گیری را کارآمدتر می‌سازد. همان‌طور که در نتایج ارائه‌شده در شکل (۵) مشاهده می‌شود، مدل پیشنهادی در تمامی مقیاس‌های پایگاه‌داده، عملکرد بهتری از نظر دقت در مقایسه با الگوریتم مبتنی بر ضرب داخلی داشته است. این امر نشان‌دهنده توانایی روش پیشنهادی در کاهش انتخاب‌های نامرتبط و ارائه پیشنهادهای هدفمندتر و کارآمدتر است.

[1] Precision

[۲] True Positive: زمانی که مدل یک نمونه را به عنوان دارای ویژگی خاصی تشخیص می‌دهد و این تشخیص درست است.

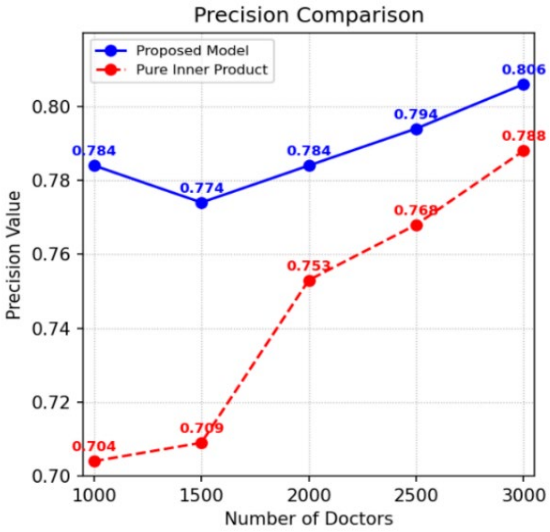
[۳] False Pasitive: زمانی که مدل به اشتباه وجود یک ویژگی را اعلام کند، درحالی‌که آن ویژگی وجود ندارد.

[4] Recall

[۵] False Negative: زمانی که مدل به اشتباه عدم وجود یک ویژگی را اعلام کند، درحالی‌که آن ویژگی وجود دارد.

[6] Efficiency

علاوه بر این، به دلیل بهره‌گیری از جست‌وجوی چند کلیدواژه‌ای روی داده‌های رمزگذاری شده، مدل ارائه‌شده حتی در شرایطی که تنوع ویژگی‌های موردنظر بیماران بالا باشد، عملکرد مطلوبی از خود نشان داده است.



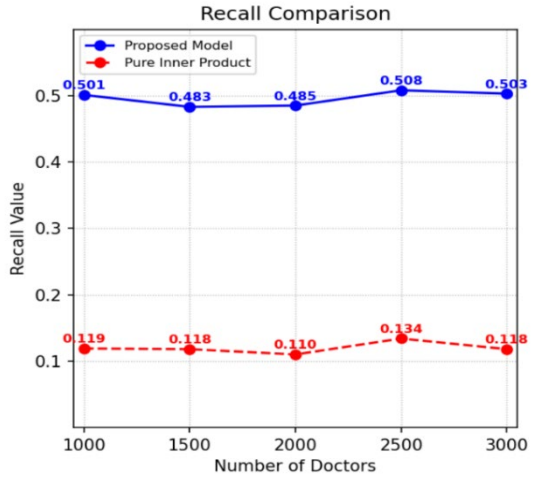
شکل (۵): مقایسه معیار دقت دو الگوریتم برای پایگاه‌های داده با اندازه‌های مختلف برای ۱۰۰ بیمار

در مجموع، بهبود قابل‌توجه دقت مدل پیشنهادی نسبت به روش ضرب داخلی بیانگر آن است که این مدل می‌تواند انتخاب‌های بهینه‌تری را در اختیار بیماران قرار داده و تجربه کاربری را ارتقا دهد، درحالی‌که هم‌زمان الزامات امنیتی و حفظ حریم خصوصی داده‌ها را نیز به طور کامل تأمین می‌کند.

مقایسه نتایج معیار بازیابی: مقدار بالاتر این معیار نشان‌دهنده توانایی مدل در بازیابی تمامی نتایج مرتبط است. به‌عبارت‌دیگر، در مسئله مورد بحث در این مقاله، پایین‌بودن این معیار بیانگر آن است که تعداد قابل‌توجهی از پزشکان مناسب در خروجی مدل لحاظ نشده‌اند که این امر می‌تواند منجر به ارائه پیشنهادهای غیربهینه به بیماران شود.

اهمیت این معیار به‌ویژه در شرایطی افزایش می‌یابد که شناسایی تمامی پزشکان مرتبط برای بیمار از اولویت بالایی برخوردار باشد و از دست‌دادن گزینه‌های مناسب هزینه بالایی به همراه داشته باشد.

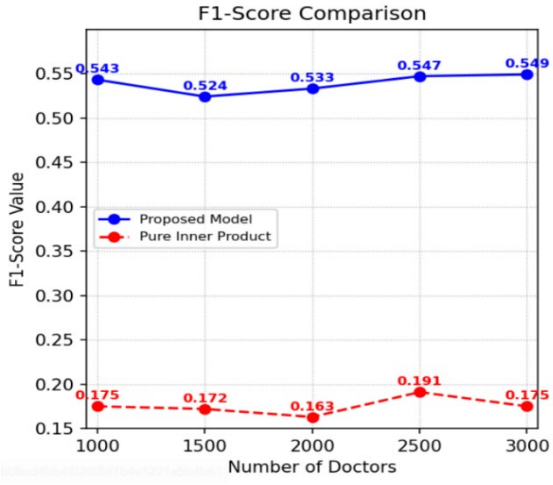
این مسئله در موقعیت‌های اورژانسی یا در مواردی که بیماران با تعداد محدودی از گزینه‌ها روبه‌رو هستند، از اهمیت بیشتری برخوردار است، زیرا عدم ارائه پیشنهادهای جامع می‌تواند تأثیر مستقیمی بر کیفیت و سرعت دریافت خدمات درمانی داشته باشد.



شکل (۶): مقایسه معیار بازیابی دو الگوریتم برای پایگاه‌های داده با اندازه‌های مختلف برای ۱۰۰ بیمار

همان‌طور که در شکل (۶) نمودار مقایسه‌ای معیار ارزیابی بازیابی بین الگوریتم پیشنهادی و الگوریتم ضرب داخلی سنتی مشاهده می‌شود، مدل ارائه‌شده در تمامی اندازه‌های پایگاه‌داده عملکرد بهتری از خود نشان داده است. این بهبود عملکرد بیانگر توانایی بالاتر مدل پیشنهادی در بازیابی جامع‌تر نتایج مرتبط و کاهش خطاهای ناشی از حذف پزشکان مناسب است. در نتیجه، بیماران می‌توانند به گزینه‌های بیشتری دسترسی داشته باشند و فرایند انتخاب پزشک بادقت و کارایی بالاتری انجام شود.

مقایسه نتایج معیار F1-score: به‌منظور ایجاد تعادل بین معیارهای دقت و بازیابی، از معیار F1-score استفاده می‌شود. شکل (۷) میانگین این معیار را برای پایگاه‌های داده با اندازه‌های مختلف، پس از پنج مرتبه شبیه‌سازی، نمایش می‌دهد. همان‌طور که انتظار می‌رفت، الگوریتم پیشنهادی در مقایسه با الگوریتم ضرب داخلی، عملکرد بهتری از نظر معیار F1 ارائه داده است. این معیار، با درنظرگرفتن هم‌زمان دقت و بازیابی، ارزیابی جامعی از کیفیت مدل ارائه می‌دهد.

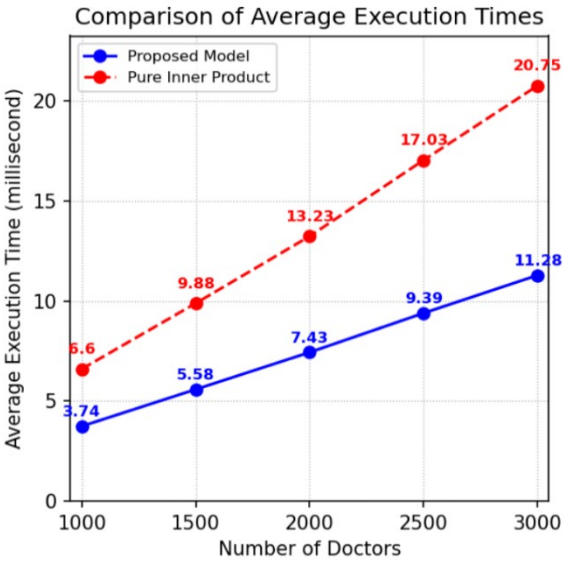


شکل (۷): مقایسه معیار $F1$ دو الگوریتم برای پایگاه‌های داده با اندازه‌های مختلف برای ۱۰۰ بیمار

به‌طورکلی، اگر با افزایش تعداد پزشکان مقدار F1 ثابت باقی بماند یا بهبود یابد، این موضوع نشان‌دهنده پایداری و کارایی بالای مدل است. در مدل پیشنهادی، تغییرات این معیار با افزایش اندازه پایگاه‌داده، اندک بوده و کاهش ناچیزی مشاهده می‌شود. بااین‌حال، این تغییرات جزئی بوده و تأثیر قابل‌توجهی بر عملکرد کلی مدل ندارد که نشان‌دهنده مقیاس‌پذیری و کارآمدی روش ارائه‌شده در شرایط مختلف است.

ازآنجاکه معیار F1 میانگین هارمونیک دقت و بازیابی است، تعادلی بین این دو معیار برقرار می‌کند. این ویژگی به‌ویژه در شرایطی که هر دو معیار اهمیت بالایی داشته و امکان اولویت‌بندی یکی بر دیگری وجود نداشته باشد، حائز اهمیت است. در واقع، F1-score به‌عنوان یک شاخص جامع و متوازن، عملکرد مدل را در سناریوهایی که هم دقت و هم بازیابی نقش کلیدی ایفا می‌کنند، به طور مؤثری ارزیابی می‌نماید.

مقایسه نتایج معیار زمان جست‌وجو: در نهایت، زمان جست‌وجوی دو الگوریتم برای پایگاه‌های داده با اندازه‌های مختلف مورد مقایسه قرار گرفته است.



شکل (۸): مقایسه زمان جست‌وجو دو الگوریتم برای پایگاه‌های داده با اندازه‌های مختلف برای ۱۰۰ بیمار

همان‌طور که در نتایج شکل (۸) مشاهده می‌شود، الگوریتم پیشنهادی در مقایسه با الگوریتم مرجع، در زمان کوتاه‌تری به نتایج بهینه دست می‌یابد. این مزیت به‌ویژه در پایگاه‌های داده بزرگ‌تر که شامل تعداد بیشتری از پزشکان هستند، از اهمیت بالاتری برخوردار است. نتایج به‌دست‌آمده نشان می‌دهند که با افزایش حجم پایگاه‌داده، اختلاف زمان جست‌وجو میان دو الگوریتم به طور قابل‌توجهی افزایش می‌یابد. این مسئله بیانگر کارایی بالاتر و مقیاس‌پذیری بهتر مدل پیشنهادی در مقایسه با الگوریتم مرجع است. به بیان دیگر، الگوریتم ارائه‌شده نه‌تنها در پایگاه‌های داده کوچک عملکرد بهتری دارد، بلکه با افزایش حجم داده‌ها، برتری آن در کاهش زمان جست‌وجو برجسته‌تر می‌شود.

۴-۸. پیچیدگی محاسباتی

این بخش به بررسی هزینه محاسباتی روش پیشنهاد شده می‌پردازد. نمادهای مورداستفاده در این بخش در جدول (۵) آورده شده‌اند:

جدول (۵): نمادهای مورد استفاده در بررسی هزینه محاسباتی

نماد	توضیح
n	اندازه بردار
T_V	مدت‌زمان عملگر ضرب داخلی
T_M	مدت‌زمان ضرب دو ماتریس در یکدیگر

ازآنجاکه الگوریتم‌های موجود برای بردارهای "نوع خدمت" و "امتیاز و زمان انتظار" مشابه یکدیگرند و هزینه محاسباتی هر الگوریتم متناسب با طول بردار است، لذا درصورتی‌که n (طول بردار نوع خدمت) به‌اندازه کافی بزرگ باشد که ۱۰ (طول بردار امتیاز و زمان انتظار) در مقابل آن مقدار ناچیز باشد، می‌توان در محاسبات هزینه از طول بردار امتیاز و زمان انتظار چشم‌پوشی نمود. به همین جهت در این بخش، از مقدار ۱۰ صرف‌نظر شده و هزینه محاسبات با همان مقدار n به‌دست‌آمده است.

هر یک از موجودیت‌های سیستم در هنگام انجام الگوریتم ها، هزینه محاسباتی متفاوتی را متحمل می‌شوند که در ادامه به تفکیک، آورده شده‌اند. پزشکان، بیمار و سرور، به ترتیب در الگوریتم‌های رمزگذاری نمایه، تولید دریچه، تبدیل نمایه و تبدیل دریچه، هر یک‌به‌دو ضرب یک بردار و یک ماتریس $(M_{i,1}^T\widetilde{W}_i^a, M_{i,2}^T\widetilde{W}_i^b)$، $(M_{q,1}^{-1}\widetilde{W}_j^a, M_{q,2}^{-1}\widetilde{W}_j^a)$، و $(M_1^T\widetilde{W}_i^a, M_2^T\widetilde{W}_i^b)$ و $(M_1^T\widetilde{W}_i^a, M_2^T\widetilde{W}_i^b)$ نیاز دارند. اجرای هر الگوریتم به‌صورت جداگانه، هزینه‌ای برابر $2nT_v$ دارد، به‌طوری که nT_v نشان‌دهنده زمان صرف شده برای انجام ضرب یک ماتریس در یک بردار را نشان می‌دهد.

علاوه‌برآن در ابتدای فرایند رمزگذاری نمایه و تولید دریچه، به‌منظور گسترش بردار اولیه، به ضرب‌های داخلی $|W_i'|^2 = W_i'.W_i'$ و $|W_j'|^2 = W_j'.W_j'$ نیاز است که خود هزینه‌ای برابر T_v دارند.

در نهایت، سرور در الگوریتم جست‌وجو نیاز دارد که ضرب داخلی نمایه‌ها و دریچه‌های تبدیل شده $M_1^T\widetilde{W}_i^a, M_2^T\widetilde{W}_i^b$ و $M_1^{-1}\widetilde{W}_j^a, M_2^{-1}\widetilde{W}_j^a$ را با هزینه زمانی $2T_v$ انجام دهد.

هزینه محاسباتی روش پیشنهادی در جدول زیر آورده شده است:

جدول (۶): هزینه محاسباتی روش پیشنهادی به تفکیک الگوریتم‌ها

الگوریتم‌ها	پزشک	بیمار	سرور
الگوریتم رمزگذاری نمایه	$T_v + 2nT_v$	-	-
الگوریتم تبدیل نمایه	-	-	$2nT_v$
الگوریتم تولید دریچه	-	$T_v + 2nT_v$	-
الگوریتم تبدیل دریچه	-	-	$2nT_v$
الگوریتم جست‌وجو	-	-	T_v
مجموع	$(2n + 1)T_v$	$(2n + 1)T_v$	$(4n + 1)T_v$

استفاده از عملیات زوج‌سازی[1]، توان رسانی[2] و عملگرهای چکیده‌سازی[3]، هر یک باعث افزایش هزینه محاسباتی و زمان اجرا می‌شوند [۳۰] و روش پیشنهادی بدون استفاده از عملگرهای مذکور و با استفاده از بردارهای دودویی، تنها از ضرب ماتریس در بردار و ضرب داخلی دو بردار به‌منظور تطبیق وظایف استفاده می‌کند.

۹. نتیجه‌گیری

در این مقاله، مسئله حفظ حریم خصوصی در سیستم‌های جمع‌سپاری مشاوره پزشکی با بهره‌گیری از مدل فضای برداری بررسی شده است. برای حل این مسئله، طرحی کارا و ایمن ارائه شده که از جست‌وجوی چند کلیدواژه و هم‌زمان چند کاربر پشتیبانی کرده و ویژگی‌هایی مانند محرمانگی نمایه و دریچه و عدم پیوندپذیری را فراهم می‌سازد.

برای دستیابی به این اهداف، از فن‌هایی همچون تجزیه ماتریس، باز رمزگذاری نیابتی و ASPE بهره گرفته شده تا تطبیق امن میان کاربران و پزشکان بر اساس حاصل‌ضرب داخلی امکان‌پذیر شود. همچنین، یک راهکار نوین به‌منظور پیش‌پردازش داده‌ها پیشنهاد شده است که دقت تطبیق را افزایش داده و محاسبات را بهینه می‌کند.

تحلیل‌های امنیتی نشان می‌دهند که این طرح، در برابر حملات شناخته‌شده مانند CPA و KPA مقاوم است و از افشای اطلاعات حساس جلوگیری می‌کند. بررسی مقاومت در برابر حملات CCA نیز به‌عنوان مسیر پژوهشی آینده مطرح شده است.

نتایج تجربی نشان می‌دهد که الگوریتم پیشنهادی نسبت به روش‌های سنتی، از نظر دقت، نرخ بازیابی، F1-score و زمان جست‌وجو عملکرد بهتری دارد و در مقیاس‌پذیری نیز موفق عمل می‌کند. این ویژگی‌ها، آن را گزینه‌ای مناسب برای مشاوره پزشکی از راه دور می‌سازد.

علاوه بر این، الگوریتم قابلیت تعمیم به حوزه‌هایی نظیر آموزش، مالی و خدمات عمومی را دارد و با ترکیب آن با روش‌های پیشرفته‌تر رمزنگاری یا یادگیری ماشین، می‌توان امنیت و کارایی را ارتقا داد. ارزیابی در محیط‌های واقعی نیز برای آینده پیشنهاد شده است.

۱۰. مراجع

[1] S.M. Purnaghi, M. Bayat, Y. Farjami, Z. Hatefi, and N. Hamian, "A Novel and Secure Model for Sharing Protected Health Record (PHR) Based on Blockchain and Attribute Based Encryption," Electronic and Cyber Defense, vol. 8, no. 1, pp. 101-124, 2020. DOR: https://dor.isc.ac/dor/20.1001.1.23224347.1399.8.1.9.9 (in Persian).

[2] R. Niemelä, M. Pikkarainen, M. Ervasti, and J. Reponen, "The change of pediatric surgery practice due to the emergence of connected health technologies," Technological Forecasting & Social Change, vol. 146, pp. 352-365, 2019. https://doi.org/10.1016/j.techfore.2019.06.001.

[3] H. Xia, and B. McKernan, "Privacy in Crowdsourcing: a Review of the Threats and Challenges," Computer Supported Cooperative Work (CSCW), vol. 29, pp. 263-301, 2020. https://doi.org/10.1007/s10606-020-09374-0.

[4] X. Fu, L. T. Yang, J. Li, X. Yang, and Z. Yang, "A Searchable Symmetric Encryption-Based Privacy Protection Scheme for Cloud-Assisted Mobile Crowdsourcing," Internet of Things, vol. 11, no. 2, pp. 1910-1924, 2024. https://doi.org/ 10.1109/JIOT.2023.3320666.

[5] B. Guo, Y. Liu, L. Wang, V.O.K. Li, J.C.K. Lam, and Z. Yu, "Task Allocation in Spatial Crowdsourcing: Current State and Future Directions," IEEE Internet of Things Journal, vol. 5, no. 3, pp. 1749-1764, 2018. https://doi.org/ 10.1109/JIOT.2018.2815982.

[6] J. Shu, and X. Jia, "Secure Task Recommendation in Crowdsourcing," IEEE Global Communications Conference (GLOBECOM), pp. 1-6, 2016. https://doi.org/ 10.1109/GLOCOM.2016.7842254.

[7] S.K. Eizadi, M. Rafiei Korkvandi, and A. Khosh Sefat, "A novel architecture for database outsourcing in cloud computing

[1] *Pairing*
[2] Exponentiation
[3] Hash

with regard to data life cycle," Electronic and Cyber Defense, vol. 2, no. 4, pp. 41-54, 2015. DOR: 20.1001.1.23224347.1393.2.4.19.3 (in Persian).

[8] S. R. B. Gummidi, "A Survey of Spatial Crowdsourcing," ACM Transactions on Database Systems, vol. 44, no. 2, pp. 1-46, 2019. https://doi.org/ 10.1145/3291933.

[9] S. Dishman, and V.G. Duffy, "The Reaches of Crowdsourcing: A Systematic Literature Review," International Conference on Human-Computer Interaction, pp. 229-248, 2021. https://doi.org/ 10.1007/978-3-030-90238-4_17.

[10] R. Lian, Y. Zheng, and C. Wang "PrivRo: A Privacy-Preserving Crowdsourcing Service with Robust Quality Awareness," Transactions on Services Computing, vol. 17, no. 4, pp. 1682-1697, 2024. https://doi.org/ 10.1109/TSC.2024.3377158.

[11] J. Shu, X. Liu, Y. Zhang, X. Jia, and R. H. Deng, "Dual-side privacy-preserving task matching for spatial crowdsourcing," Journal of Network and Computer Applications, vol. 123, pp. 101-111, 2018. https://doi.org/ 10.1016/j.jnca.2018.09.007.

[12] J. Shu, X. Jia, K. Yang, and H. Wang, "Privacy-Preserving Task Recommendation Services for Crowdsourcing," Transactions on Services Computing, vol. 14, no. 1, pp. 235-247, 2018. https://doi.org/ 10.1109/TSC.2018.2791601.

[13] T. Peng, W. Zhong, G. Wang, S. Zhang, E. Luo, and T. Wang, "Spatiotemporal-aware Privacy-preserving Task Matching in Mobile Crowdsensing," Internet of Things Journal, vol. 11, no. 2, pp. 2394-2406, 2023. https://doi.org/ 10.1109/JIOT.2023.3292284.

[14] Y. Gong, Y. Guo, and Y. Fang, "A privacy-preserving task recommendation framework for mobile crowdsourcing," IEEE Global Communications Conference, pp. 588-593, 2014. https://doi.org/ 10.1109/GLOCOM.2014.7036871.

[15] R. Zhang, R. Xue, and L. Liu, "Searchable Encryption for Healthcare Clouds: A Survey," Transactions on Services Computing, vol. 11, no. 6, pp. 978-96, 2017. https://doi.org/10.1109/TSC.2017.2762296.

[16] Y. Watanabe, T. Nakai, K. Ohara, T. Nojima, Y. Liu, M. Iwamoto, and K. Ohta, "How to Make a Secure Index for Searchable Symmetric Encryption, Revisited," IEICE Transactions on Fundamentals of Electronics, vol. 105, no. 12, pp. 1559-1577, 2022. https://doi.org/ 10.1587/transfun.2021EAP1163.

[17] Q. Liu, Y. Guo, J. Wu, and G. Wang, "Effective Query Grouping Strategy in Clouds," Computer Science and Technology, vol. 32, no. 6, pp. 1231-1249, 2017. https://doi.org/ 10.1007/s11390-017-1797-9.

[18] J. Shu, X. Liu, X. Jia, and K. Yan, "Anonymous Privacy-Preserving Task Matching in Crowdsourcing," Internet of Things Journal, vol. 5, no. 4, pp. 3068-3078, 2018. https://doi.org/ 10.1109/JIOT.2018.2830784.

[19] R. A. Popa, and N. Zeldovich, "Multi-Key Searchable Encryption," IACR Cryptology ePrint Archive, 2013.

[20] H. Bao, Z. Wang, R. Lu, C. Huang, and B. Li, "TAMT: Privacy-Preserving Task Assignment with Multi-Threshold Range Search for Spatial Crowdsourcing Applications," IEEE Transactions on Big Data, pp. 1-13, 2024. https://doi.org/ 10.1109/TBDATA.2024.3403374.

[21] A. Kiayias, O. Oksuz, A. Russell, Q. Tang, and B. Wang, "Efficient Encrypted Keyword Search for Multi-user Data Sharing," European symposium on research in computer security, pp. 173-195, 2016. https://doi.org/ 10.1007/978-3-319-45744-4_9.

[22] F. Zhao, T. Nishide, and K. Sakurai, "Multi-User Keyword Search Scheme for Secure Data Sharing with Fine-Grained Access Control," International Conference on Information Security and Cryptology-ICISC 2011: 14th International Conference, Seoul, Korea, pp. 406-418, 2012. https://doi.org/ 10.1007/978-3-642-31912-9_27.

[23] K. Liang, C. Su, J. Chen, and J. K. Liu, "Efficient Multi-Function Data Sharing and Searching Mechanism for Cloud-Based Encrypted Data," 11th ACM on Asia Conference on Computer and Communications Security, pp. 83-94, 2016. https://doi.org/ 10.1145/2897845.289786.

[24] H. Li, D. Liu, Y. Dai, T. H. Luan, and X. S. Shen, "Enabling Efficient Multi-Keyword Ranked Search Over Encrypted Mobile Cloud Data Through Blind Storage," Emerging Topics in Computing, vol. 3, no. 1, pp. 127-138, 2014. https://doi.org/ 10.1109/TETC.2014.2371239.

[25] Y. Miao, W. Zheng, X. Jia, X. Liu, K. K. R. Choo, and R. H. Deng, "Ranked Keyword Search over Encrypted Cloud Data Through Machine Learning Method," IEEE Transactions on Services Computing, vol. 16, no. 1, pp. 525-536, 2021. https://doi.org/ 10.1109/TSC.2021.3140098.

[26] N. Cao, C. Wang, M. Li, and K. Ren, "Privacy-preserving multi-keyword ranked search over encrypted cloud data," Transactions on parallel and distributed systems, vol. 25, no. 1, pp. 222-233, 2014. https://doi.org/ 10.1109/TPDS.2013.45.

[27] W. K. Wong, D. W. Cheung, B. Kao, and N. Mamoulis, "Secure kNN Computation on Encrypted Databases," International

Conference on Management of data, pp. 139-152, 2009. https://doi.org/10.1145/1559845.155986.

[28] M. Zhang, Y. Chen, and J. Huang, "SE-PPFM: A Searchable Encryption Scheme Supporting Privacy-Preserving Fuzzy Multikeyword in Cloud Systems," Systems Journal, vol. 15, no. 2, pp. 2980-2988, 2020. https://doi.org/10.1109/JSYST.2020.2997932.

[29] F. Song, Z. Qin, D. Liu, J. Zhang, X. Lin, and X. Shen, "Privacy-Preserving Task Matching with Threshold Similarity Search via Vehicular Crowdsourcing," Transactions on Vehicular Technology, vol. 70, no. 7, pp. 7161-7175, 2021. https://doi.org/10.1109/TVT.2021.3088869.

[30] G. Soltan, A. Wong, and C. S. Yang, "A Vector Space Model for Automatic Indexing," Information Retrieval and Language Processing, vol. 18, no. 11, pp. 613-620, 1975. https://doi.org/10.1145/361219.361220.