



بررسی عملکرد الگوریتم ژنتیک کوانتومی در مسائل زیست‌شناسی

معین عطاری^۱، نیره مجد^{۲*}

*نویسنده مسئول، دریافت: ۱۴۰۳/۰۵/۰۳، بازنگری: ۱۴۰۳/۰۸/۲۷، پذیرش: ۱۴۰۳/۰۹/۰۸ (تاریخ‌ها توسط نشریه وارد می‌شوند)

^۱ فارغ‌التحصیل کارشناسی ارشد، الگوریتم و محاسبات، دانشکده علوم مهندسی، پردیس دانشکده‌های فنی، دانشگاه تهران، تهران
^۲ استادیار، الگوریتم و محاسبات، دانشکده علوم مهندسی، پردیس دانشکده‌های فنی، دانشگاه تهران، تهران

چکیده

در این پژوهش، یک مسئله از علم زیست‌شناسی، با عنوان پیش‌بینی ساختار پروتئینی (PSP)، مورد بررسی قرار گرفته است. مسئله PSP در علم زیست‌شناسی به‌عنوان مسئله محاسباتی و مهم شناخته می‌شود که کاربردهای زیادی مانند درمان سرطان و طراحی دارو دارد، و در صنایع مختلف استفاده می‌شود. مسئله PSP از نوع مسائل NP-Hard است. در این پژوهش، با استفاده از الگوریتم ژنتیک کوانتومی، که بر اساس محاسبات کوانتومی و الگوریتم ژنتیک کلاسیک طراحی شده است، توانستیم بهترین جواب بهینه سراسری مسئله PSP را در زمان قابل قبول بهتری نسبت به الگوریتم‌های مشابه کلاسیک پیدا نماییم. از آنجاکه در این پژوهش هدف ما حل مسئله PSP با رویکرد کوانتومی است، با توجه به درجه آزادی حرکت پروتئین و نیز محدودیت ابعاد کیوبیت‌ها و کیوبت‌ها، توانستیم کدبندی درست حرکت را با توجه به محدودیت ساختار ذاتی آن‌ها پیاده‌سازی نماییم. همچنین در این پژوهش برای الگوریتم‌های ژنتیک کوانتومی مبتنی بر کیوبت، یک جدول جستجوی جدید برای بروز رسانی ژن‌های کیوبت‌ها معرفی شده است، که از سه گیت چرخشی برای بروز رسانی ژن‌های کیوبت‌ها استفاده می‌کند. این جدول جستجوی جدید با سایر روش‌های موجود مقایسه شده است که، دارای تنوع و سرعت بهتر است.

کلمات کلیدی: الگوریتم ژنتیک، محاسبات کوانتومی، الگوریتم ژنتیک کوانتومی، پیش‌بینی ساختار پروتئینی

۱- مقدمه

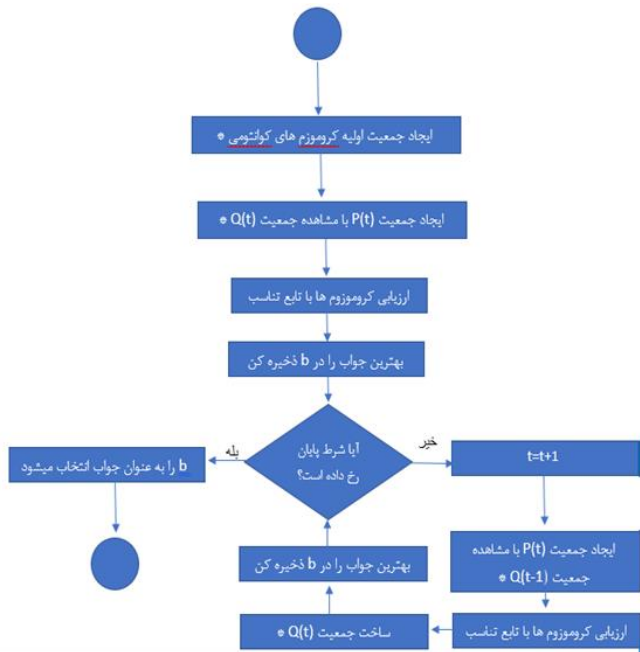
یکی از مهم‌ترین و چالشی‌ترین مسائل در علم زیست‌شناسی، مسئله PSP است. مسئله PSP در بسیاری از زمینه‌های بیولوژیکی مانند طراحی دارو [۱]، بررسی داده‌های تشدید مغناطیسی هسته‌ای [۲] و درمان از بیماری‌ها مثل آلزایمر، سرطان و پارکینسون [۳،۴] کاربرد دارد. در مسئله PSP ارتباط بین دنباله‌های آمینواسید و چیدمان پروتئین مورد مطالعه قرار می‌گیرد. جواب بهینه مسئله PSP، می‌تواند مشکلات چالش برانگیز و جدی شامل بیوشیمی، بیوفیزیک، زیست‌شناسی ساختاری، زیست‌شناسی مولکولی و زیست‌شناسی محاسباتی را حل کند.

بسیاری از محققان، الگوریتم‌هایی را برای یافتن جواب مسئله PSP بر اساس فرضیه ترمودینامیکی پیشنهاد کرده‌اند، زیرا دارای کمترین انرژی آزاد گیبس^۲ در شرایط فیزیولوژیکی برای ساختارهای پروتئینی است [۵]. مسئله PSP را می‌توان به دو مسئله فرعی تقسیم کرد: (۱) تعریف تابع انرژی مناسب جهت مقایسه جواب‌های مسئله و تشخیص جواب‌های قابل قبول از جواب‌های غیرقابل قبول (۲) معرفی الگوریتم مناسبی که دارای ابزارهای سازگار و قدرتمند برای مقابله با مشکل فضای جستجوی بزرگ است.

هر ساختار یا چیدمان این رشته پروتئین دارای ویژگی‌های متفاوت و منحصر به فرد است که با چیدمان و ساختار دیگر متفاوت است. جیمز و توفیک^۳ در سال ۲۰۰۳ تنوع ساختاری پروتئین‌ها را مورد بازبینی قرار دادند و فرضیه‌ای را بر اساس «دیدگاه جدید» پروتئین‌ها بیان کردند که به موجب آن یک رشته پروتئینی می‌تواند ساختارها و عملکردهای متعددی را اتخاذ کند [۶]. کرسنزی^۴ در سال ۱۹۹۸ بیان کرد که مسئله PSP یک مسئله NP-Hard است [۷]. وتلافی^۵ در سال ۱۹۷۳ بیان کرد که چیدمان پروتئین‌ها به کمک یک تابع جستجوی جامع حداقل ده‌ها مرتبه سریع‌تر انجام می‌شود [۸]. بوی^۶ در سال ۲۰۰۵ الگوریتم ژنتیک برای پیش‌بینی جواب بهینه سراسری ساختار دوبعدی مسئله PSP ارائه کرد [۹]. خیماسیا^۷ در سال ۱۹۹۷ الگوریتم ژنتیک برای پیش‌بینی جواب بهینه سراسری ساختار سه‌بعدی مسئله PSP ارائه کرد [۱۰]. فابیو^۸ در سال ۲۰۱۴ یک الگوریتم ژنتیک حداقل چندگانه^۹ را برای یافتن جواب بهینه سراسری مسئله PSP در مدل شبکه سه‌بعدی HP ارائه می‌کند [۱۱] که توانایی یافتن جواب بهینه سراسری رشته پروتئینی ۱۳۶ عضوی را دارد.

۲- روش تحقیق

در شکل (۱) گام‌های موجود در الگوریتم ژنتیک کوانتومی را نشان می‌دهد. مراحل شکل (۱) که دارای ستاره (*) هستند، به این معنی است که آن مراحل دارای محاسبات کوانتومی هستند، مراحل دیگر بدون ستاره دارای محاسبات کلاسیک مانند الگوریتم ژنتیک هستند. در ادامه این فصل، مراحل محاسبات کوانتومی، الگوریتم ژنتیک کوانتومی به‌طور خلاصه توضیح داده شده است.



شکل ۱- گام‌های موجود در الگوریتم ژنتیک کوانتومی [۱۱]

در الگوریتم‌های ژنتیک ارائه شده، روش‌های کدگذاری برای مسئله PSP، در ساختار دوبعدی و ساختار سه‌بعدی، هر کروموزوم دارای $N-1$ ژن است، در ساختار دوبعدی، هر ژن می‌تواند بر روی نمادهای $\{R, L, U, D\}$ و در ساختار سه‌بعدی، هر ژن را می‌توان بر روی نمادهای $\{U, R, D, L, F, B\}$ تخصیص داد. اما در الگوریتم ژنتیک کوانتومی، کیوبیت به‌عنوان واحد اطلاعات پایه، ممکن است در حالت پایه $|0\rangle$ یا حالت پایه $|1\rangle$ باشد، یا در هر برهم نهی از این دو، کیوبیت به‌عنوان واحد اطلاعات در محاسبات کوانتومی دارای سه است. می‌تواند در هر یک از سه حالت یا در هر برهم نهی از آن‌ها باشد. بنابراین، لازم است روش کدگذاری جدیدی ایجاد شود که بتواند چهار نماد حرکتی برای ساختار دوبعدی و شش نماد حرکتی برای ساختار سه‌بعدی با استفاده از کیوبیت و کیوبیت کدگذاری کند.

از دو کیوبیت برای کدگذاری چهار نماد حرکتی در ساختار دوبعدی برای الگوریتم ژنتیک کوانتومی برای مسئله PSP در ساختار دوبعدی مطابق جدول (۱) استفاده می‌کنیم. بنابراین، در هر کروموزوم کوانتومی برای ساختار دوبعدی مسئله ما $2(N-1)$ ژن کیوبیتی داریم که هر دو ژن بیانگر یک حرکت در ساختار دوبعدی است.

جدول ۱- جمعیت کدگذاری نمادهای حرکتی ساختار دوبعدی توسط الگوریتم

ژنتیک کوانتومی [۱۴]		
نام حرکت	مقدار کیوبیت اول	مقدار کیوبیت دوم
بالا (Up)	۰	۰
پایین (Down)	۱	۰
راست (Right)	۰	۱
چپ (Left)	۱	۱

در سال ۲۰۲۱ صادق، بوروی اند بومدینه، نیپلم، مسئله PSP در مدل دوبعدی مثلثی مطالعه و بررسی کردند و الگوریتم ژنتیک برای آن ارائه کردند [۱۲]. ژانگ^{۱۰} در سال ۲۰۱۰ الگوریتم جستجوی تابو ژنتیکی^{۱۱} را برای یافتن جواب بهینه سراسری مسئله PSP در مدل شبکه سه‌بعدی HP ارائه کرد [۱۳]، که ترکیب الگوریتم ژنتیک که قابلیت بالایی در جستجوی سراسری دارد و روش جستجوی تابو که در جستجوی محلی بسیار توانمند است. عطاری و مجد در سال ۲۰۲۲ یک الگوریتم ژنتیک کوانتومی برای پیش‌بینی جواب بهینه سراسری ساختار دوبعدی مسئله PSP ارائه کردند [۱۴] که این الگوریتم بر اساس الگوریتم ژنتیک است برای سرعت بخشیدن به انجام محاسبات، با محاسبات کوانتومی ترکیب شده است.

در پژوهش حاضر از پروتئین‌های شبکه^{۱۲} استفاده می‌شود زیرا آن‌ها ساده‌ترین مدل مورد استفاده برای شبیه‌سازی چیدمان پروتئین هستند [۱۵]. پروتئین واقعی یک مولکول بزرگ است که حاوی صدها یا هزاران اتم است و شبیه‌سازی رفتار یک پروتئین واقعی با فناوری فعلی غیرممکن است. پروتئین‌های شبکه عملکرد الگوریتم‌ها را تسریع می‌کنند و به دستیابی به انرژی بهینه سراسری کمک می‌کنند. پروتئین‌های شبکه از دو قسمت تشکیل شده‌اند. (i) هر اتم یک آمینواسید به‌عنوان "مه‌ره"^{۱۳} در نظر گرفته می‌شود، و (ii) همه این "مه‌ره‌ها" به یک شبکه (معمولاً مکعبی) محدود می‌شوند [۱۴].

در این پژوهش از پروتئین شبکه‌ای به نام مدل HP استفاده شده است که شامل دو نوع مه‌ره H (آب‌گریز یا غیر قطبی) و P (آب‌دوست یا قطبی) است. در مسئله PSP، دنباله آمینواسید وجود دارد و هدف یافتن چیدمان بهینه این زنجیره‌های متصل است که بهترین عملکرد انرژی (کمترین انرژی) را توسط ساختارهای دوبعدی یا ساختار سه‌بعدی دارد. مقالات زیادی وجود دارد که از الگوریتم ژنتیک برای یافتن جواب بهینه سراسری مسئله PSP استفاده می‌کنند [۱۳-۱۱]، اما تاکنون هیچ الگوریتم برای مسئله PSP معرفی نشده که از محاسبات کوانتومی استفاده کند. با توجه به اینکه مسئله PSP، که یک مسئله بهینه‌سازی از نوع NP-Hard است، استفاده از محاسبات کوانتومی باعث افزایش سرعت انجام محاسبات می‌شود، و ترکیب آن با الگوریتم ژنتیک به‌عنوان الگوریتم تکاملی معروف و قدرتمند باعث یافتن جواب بهینه سراسری را در زمان کمتری می‌شود.

امروز نیز از روش‌ها و الگوریتم‌های نوینی برای مسئله PSP مانند یادگیری ماشین و یادگیری عمیق استفاده می‌شود. AlphaFold2 یک مدل یادگیری عمیق است که توسط شرکت DeepMind توسعه یافته و توانایی شگفت‌انگیزی در پیش‌بینی ساختار سه‌بعدی پروتئین‌ها از روی توالی آمینواسیدی آن‌ها دارد. این مدل، که به‌عنوان یکی از بزرگ‌ترین پیشرفت‌ها در زیست‌شناسی محاسباتی در دهه‌های اخیر شناخته می‌شود، به‌طور چشمگیری زمان و هزینه مورد نیاز برای تعیین ساختار پروتئین‌ها را کاهش داده است [۱۶-۱۵]. البته در این مقاله تنها روش‌های کلاسیک ژنتیک موجود برای مسئله PSP مورد بررسی و ارزیابی می‌شود.

HelixFold-Single ابتدا یک مدل زبان پروتئینی در مقیاس بزرگ را با هزاران میلیون ساختار اولیه با استفاده از الگوی یادگیری خود نظارتی پیش‌آموز می‌کند، که به‌عنوان جایگزینی برای MSAها برای یادگیری اطلاعات تکاملی مشترک استفاده خواهد شد [۱۷].

در این پژوهش الگوریتم جدیدی از نوع الگوریتم ژنتیک کوانتومی برای مسئله PSP معرفی شده است که می‌تواند جواب بهینه سراسری مسئله در ساختار دوبعدی سریع‌تر، ولی با حفظ دقت، پیدا کند. همچنین می‌تواند جواب بهینه سراسری برای مسئله PSP در ساختار سه‌بعدی پیدا کند علاوه بر معرفی الگوریتم ژنتیک کوانتومی جدید برای مسئله PSP در ساختار دوبعدی و سه‌بعدی، یک جدول جستجوی^{۱۴} جدید برای تولید جمعیت جدید کوانتومی نیز معرفی می‌شود که در مقایسه با سایر روش‌های موجود برای تولید جمعیت جدید کوانتومی سریع‌تر و دارای تنوع بیشتری است.

عملگر	$f(X) \geq f(B)$	b_i	x_i
Increase a	False	0	0
Increase b	False	1	0
Increase c	False	2	0
Increase a	False	0	1
Increase b	False	1	1
Increase c	False	2	1
Increase a	False	0	2
Increase b	False	1	2
Increase c	False	2	2
Increase a	True	0	0
Increase a	True	1	0
Increase a	True	2	0
Increase b	True	0	1
Increase b	True	1	1
Increase b	True	2	1
Increase c	True	0	2
Increase c	True	1	2
Increase c	True	2	2

از یک کیوبیت و یک کیوتریت برای کدگذاری شش نماد حرکتی در الگوریتم ژنتیک کوانتومی برای مسئله PSP در ساختار سه بعدی مطابق جدول (۲) استفاده می‌کنیم. بنابراین، ما $2(N-1)$ ژن داریم که، $N-1$ ژن آن، شامل ژن‌های کیوبیت (ژن‌های فرد) و $N-1$ ژن آن، شامل ژن‌های کیوتریت (ژن‌های زوج) است. که هر دو ژن کیوبیت و کیوتریت بیانگر یک حرکت در ساختار سه بعدی است.

جدول ۲- جمعیت کدگذاری نمادهای حرکتی ساختار سه بعدی توسط الگوریتم ژنتیک کوانتومی

نام حرکت	مقدار کیوبیت	مقدار کیوتریت
بالا (Up)	۰	۰
پایین (Down)	۱	۱
راست (Right)	۰	۲
چپ (Left)	۱	۰
جلو (Front)	۰	۱
عقب (Behind)	۱	۲

جدول (۳) جدول جستجوی ژن‌های کیوتریت را نشان می‌دهد که برای به‌روزرسانی ژن‌های کیوتریت به یک جدول جستجوی مشترک نیاز دارد. جدول (۴) جدول جستجوی کمکی^{۱۵} برای ژن‌های کیوبیت و شکل (۲) گیت‌های چرخش این جدول جستجو را نشان می‌دهد.

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{bmatrix}$$

$$R_y(\beta) = \begin{bmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{bmatrix}$$

$$R_z(\gamma) = \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

شکل ۲- گیت‌های چرخش جدول جستجوی ژن‌های کیوتریت

جدول جستجو برای ژن‌های کیوبیت فقط یک گیت کوانتومی دارد. جدول (۵) جدول جستجوی ژن‌های کیوبیت را نشان می‌دهد. شکل (۳) گیت چرخشی این جدول جستجو را نشان می‌دهد. همان‌طور که در شکل (۳) نشان داده شده است، دامنه‌های یک کیوبیت را می‌توان در فضای دوبعدی توسط یک دایره نشان داد. در این شکل مقدار α با x و مقدار β با y نشان داده شده است. اگر مقدار ضریب دامنه α با β ($sign(\alpha\beta_i)$) مثبت شود یعنی در ناحیه ۱ و ۳ شکل (۴) هستیم و اگر منفی شود یعنی در ناحیه ۲ و ۴ شکل (۴) هستیم.

جدول ۵- جستجو برای ژن‌های کیوبیت

β_i	$s(\alpha_i, \beta_i)$			$\Delta\theta$	$f(X) \geq f(B)$	b_i	x_i
	$\alpha_i = 0$	$\alpha_i \beta_i < 0$	$\alpha_i \beta_i > 0$				
0	0	0	0	0	False	0	0
0	0	0	0	0	True	0	0
0	0	0	0	0	False	1	0
0	$1\pm$	+1	-1	0.05	True	1	0
0	$1\pm$	+1	-1	0.01	False	0	1
$1\pm$	0	-1	+1	0.025	True	0	1
$1\pm$	0	-1	+1	0.005	False	1	1
$1\pm$	0	-1	+1	0.025	True	1	1

جدول ۳- جستجو برای ژن‌های کیوتریت

افزایش γ	افزایش β	افزایش α	Sign(amp)			Octan
			γ	β	α	
Rotation X (+)	Rotation X (-)	Rotation Y (+)	+	+	+	۱
Rotation Y (-)	Rotation Z (+)	Rotation Z (-)	+	+	+	۱
Rotation X (-)	Rotation X (+)	Rotation Y (-)	-	+	+	۲
Rotation Y (+)	Rotation Z (+)	Rotation Z (-)	-	+	+	۲
Rotation X (-)	Rotation X (+)	Rotation Y (+)	+	-	+	۳
Rotation Y (-)	Rotation Z (-)	Rotation Z (+)	+	-	+	۳
Rotation X (+)	Rotation X (-)	Rotation Y (-)	-	-	+	۴
Rotation Y (+)	Rotation Z (-)	Rotation Z (+)	-	-	+	۴
Rotation X (+)	Rotation X (-)	Rotation Y (-)	+	+	-	۵
Rotation Y (-)	Rotation Z (+)	Rotation Z (-)	+	+	-	۵
Rotation X (-)	Rotation X (+)	Rotation Y (+)	-	+	-	۶
Rotation Y (+)	Rotation Z (-)	Rotation Z (+)	-	+	-	۶
Rotation X (-)	Rotation X (+)	Rotation Y (-)	+	-	-	۷
Rotation Y (+)	Rotation Z (+)	Rotation Z (-)	+	-	-	۷
Rotation X (+)	Rotation X (-)	Rotation Y (+)	-	-	-	۸
Rotation Y (-)	Rotation Z (+)	Rotation Z (-)	-	-	-	۸

جدول ۴- جستجو کمکی برای ژن‌های کیوتریت

$$F_2 = [1 + (x_1 + x_2 + 1)^2 \times (19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)] \times [30 + (2x_1 - 3x_2)^2 \times (18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)]$$

$$-2 \leq x_i \leq 2, i = 1,2 \quad (۳)$$

تابع آزمایش فوق تنها یک جواب کمینه سراسری دارد

$$F_2(0, -1) = 3 \quad (۴)$$

Schafer Function

$$F_3 = 0.5 + \frac{\sin^2(x_1^2 + x_2^2) - 0.5}{[1.0 + 0.001(x_1^2 + x_2^2)]^2}$$

$$-100 \leq x_i \leq 100, i = 1,2 \quad (۵)$$

تابع آزمایش فوق تنها یک جواب کمینه سراسری دارد

$$F_3(0,0) = 0 \quad (۶)$$

Monopole and Six-Peak Camelback Function

$$F_4 = 10 + \frac{\sin(\frac{1}{x})}{0.1 + (x - 0.16)^2}$$

$$0 \leq x \leq 1 \quad (۷)$$

تابع آزمایش فوق تنها یک جواب کمینه سراسری دارد

$$F_4(0.1275) = 19.8949 \quad (۸)$$

Dual-Pole and Six-Peak Function

$$F_5 = \left(4 - 21x_1^2 + \frac{1}{3}x_1^4\right)x_1^2 + x_1x_2 + (-4 + 4x_2^2)x_2^2 \quad (۹)$$

$$-3 \leq x_i \leq 3, i = 1,2$$

تابع آزمایش فوق دو جواب کمینه سراسری دارد

$$F_5(-0.0898, 0.7126) = F_5(0.0898, -0.7126) = -1.031628 \quad (۱۰)$$

Multipeak Positive Function

$$F_6 = e^{-0.001x} \cos^2(0.8x)$$

$$-2 \leq x \leq 5 \quad (۱۱)$$

تابع آزمایش فوق تنها یک جواب کمینه سراسری دارد

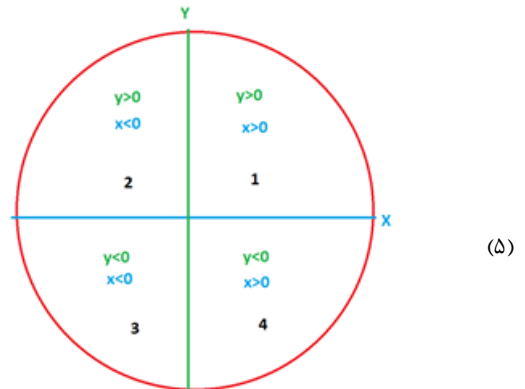
$$F_6(0) = 1 \quad (۱۲)$$

Ackley Function

$$F_7 = -20e^{-0.2\sqrt{\frac{1}{n}\sum_{i=1}^n x_i^2}} - e^{\left(\frac{1}{n}\sum_{i=1}^n \cos(2\pi x_i)\right)} + 22.71828 \quad (۱۳)$$

$$U(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

شکل ۳- گیت چرخش جدول جستجوی ژنهای کیوبیت



شکل ۴- نمایش دامنه‌های کیوبیت در فضای ۲ بعدی توسط دایره

جدول ۶- فهرست نمادها

نماد	توضیح
α	احتمال قرار داشتن در حالت $\{0\}$
β	احتمال قرار داشتن در حالت $\{1\}$
γ	احتمال قرار داشتن در حالت $\{2\}$
b_i	ژن ام در بهترین کروموزوم کلاسیک یافته شده
x_i	ژن ام در کروموزوم کلاسیک در حال ارزیابی
SD	انحراف استاندارد تابع تناسب
\hat{t}	مقدار متوسط زمان اجرا
\hat{f}	مقدار متوسط بهترین تابع تناسب یافته شده
f_{opt}	بهترین تابع تناسب یافته شده
T	تعداد نسل‌های ایجادشده برای بررسی عملکرد یکبار اجرای الگوریتم
R	تعداد تکرار الگوریتم
S	تعداد کروموزوم در هر نسل

۳- جدول جستجوی برای کیوتریت

برای بررسی عملکرد و توانایی جدول جستجوی برای بروز رسانی ژنهای کیوتریت از هشت تابع آزمایش استاندارد [۲۰-۲۱] استفاده می‌شود. از این هشت تابع برای ارزیابی و بررسی عملکرد روش‌های بروز رسانی ژنهای کوانتومی در الگوریتم ژنتیک کوانتومی استفاده می‌شود. این هشت تابع استاندارد به شرح ذیل است.

De Jong Function

$$F_1 = 100(x_1^2 - x_2^2) + (1 - x_1)^2$$

$$-2.048 \leq x_i \leq 2.48, i = 1,2 \quad (۱)$$

تابع آزمایش فوق تنها یک جواب کمینه سراسری دارد

$$F_1(1,1) = 0 \quad (۲)$$

Coldstein Price Function

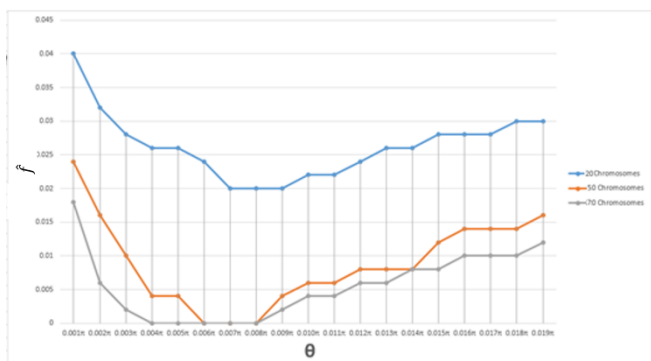
$(HP)_2PH(HP)_2(PH)_2HP(PH)_2$	۲۰	۱
$H_2P_2(HP_2)_6H_2$	۲۴	۲
$P_2HP_2(H_2P_4)_3H_2$	۲۵	۳
$P(P_2H_2)_2P_5H_5(H_2P_2)_2P_2H(HP_2)_2$	۳۶	۴
$P_2H(P_2H_2)_2P_5H_{10}P_6(H_2P_2)_2HP_2H_5$	۴۸	۵

عملکرد الگوریتم ژنتیک کوانتومی به دو پارامتر، اندازه جمعیت و زوایای چرخش جدول جستجو بستگی دارد. در این پژوهش از دو نوع جدول جستجو استفاده شد، یک جدول جستجوی برای بروز رسانی ژن‌های کیوبیت و یک جدول جستجو برای بروز رسانی ژن‌های کیوبیت استفاده می‌شود.

در ابتدا، برای بررسی عملکرد الگوریتم شرط خاتمه داده شده است (ایجاد $T=100$ نسل توسط الگوریتم). سپس الگوریتم را اجرا کرده و در نهایت مقدار به دست آمده از تابع تناسب مربوطه را به عنوان خروجی نمایش می‌دهد. در شکل ۱۹، f به معنای میانگین خروجی‌ها برای ده اجرای الگوریتم ($R=10$) برای زاویه چرخش شناخته شده و اندازه جمعیت شناخته شده است.

از آنجایی که این شبیه‌سازی برای بررسی زاویه چرخش کویت‌ها در جدول جستجو و اندازه جمعیت (ها) روی عملکرد QGA برای یافتن بهترین راه‌حل سراسری بهینه برای مسئله PSP است، بنابراین ما زوایای چرخش مختلف و جمعیت‌های مختلف را بررسی کرده‌ایم. اندازه ما مقادیر زوایای چرخش را بین $\pi \cdot 0.001$ تا $\pi \cdot 0.19$ با مراحل 0.001 تغییر دادیم (x ستون‌ها در شکل ۵). برای هر زاویه چرخش، الگوریتم را ده بار اجرا کردیم ($R=10$) و میانگین خروجی ده اجرا را ثبت کردیم (ستون \bar{y} در شکل ۵).

این شبیه‌سازی ابتدا با ۲۰ کروموزوم در هر جمعیت (خط آبی در شکل ۵)، سپس ۵۰ کروموزوم در هر جمعیت (خط نارنجی در شکل ۵) و در نهایت ۷۰ کروموزوم در هر جمعیت (خط خاکستری در شکل ۵) انجام شد.



شکل ۵- تأثیر زاویه چرخش کویت‌ها در جدول جستجو و اندازه جمعیت بر میانگین تابع تناسب اولین دنباله جدول ۸ در ساختار سه‌بعدی برای $R=10$.

همان‌طور که از شکل ۵ می‌توان دید، جمعیت کمتر نمی‌تواند راه‌حل بهینه جهانی مسئله را پیدا کند و ممکن است در راه‌حل بهینه محلی قرار گیرد. اندازه جمعیت بزرگ می‌تواند راه‌حل بهینه سراسری را پیدا کند، اما زمان بسیار بیشتری برای تکمیل محاسبات صرف می‌شود. بنابراین، توصیه می‌کنیم اندازه جمعیت را ۵۰ تنظیم کنید. همان‌طور که در شکل ۵ مشاهده می‌کنید، با در نظر گرفتن زوایای چرخش بین $\pi \cdot 0.01$ و $\pi \cdot 0.12$ ، می‌توان راه‌حل بهینه‌تری را پیدا کرد.

در این قسمت تأثیر زوایای چرخش بر مقادیر انحراف معیار (SD) برای دومین دنباله جدول ۸ بررسی شده است. بدین منظور در شبیه‌سازی تعداد جمعیت‌ها در هر زمان اجرا برابر با ۵۰ عدد ($T=50$) و تعداد کروموزوم‌ها در هر جمعیت برابر با ۲۰ عدد ($S=20$) است. در این شبیه‌سازی، مقادیر زاویه چرخش θ

$$-5 \leq x_i \leq 5, i = 1, 2 \quad (14)$$

تابع آزمایش فوق تنها یک جواب کمینه سراسری دارد

$$F_7(0,0) = 0$$

Rastrigin Function

$$F_8 = 10n + \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i)] \quad (15)$$

$$-5.12 \leq x_i \leq 5.12, i = 1, 2$$

تابع آزمایش فوق تنها یک جواب کمینه سراسری دارد

$$F_8(0,0) = 0 \quad (16)$$

ما ۸ تابع آزمایش ذکر شده در این بخش را با اندازه جمعیت برابر با ۱۰ تا ۳۰ و زاویه چرخش برابر با $\pi \cdot 0.007$ تا $\pi \cdot 0.09$ شبیه‌سازی کردیم. جدول (۷) نتایج ما را نشان می‌دهد. در جدول (۷)، f_{opt} به معنای مقدار بهینه تابع برازندگی است. \hat{f} مقدار متوسط تابع برازندگی بهترین جواب یافته شده است. sd انحراف استاندارد تابع برازندگی است. \hat{t} مقدار متوسط زمان سپری شده مورد نیاز برای اجرای الگوریتم است. نتایج به دست آمده با مرجع شماره [۲۰] مقایسه شده است.

همچنین در جدول (۷) ستون QGA، نتایج به دست آمده از الگوریتم ژنتیک کوانتومی بر پایه کیوبیت است. ستون QGA_3، نتایج به دست آمده از الگوریتم ژنتیک کوانتومی بر پایه کیوبیت با روش بروز رسانی ژن‌ها بدون استفاده از جدول است. ستون QGA_Lookup_Table، الگوریتم ژنتیک کوانتومی جدید ارائه شده بر پایه کیوبیت با روش بروز رسانی ژن‌ها با استفاده از جدول است.

جدول ۷- مقایسه جدول جستجوی ارائه شده با الگوریتم‌های دیگر

Function	QGA				QGA_3				QGA_Lookup_Table			
	f_{opt}	\hat{f}	SD	\hat{t}	f_{opt}	\hat{f}	SD	\hat{t}	f_{opt}	\hat{f}	SD	\hat{t}
F_1	0.000	0.011	0.011	0.035	0.000	0.005	0.002	0.024	0.000	0.000	0.001	0.021
F_2	3.000	3.050	0.132	0.034	3.000	3.000	0.001	0.022	3.000	3.000	0.001	0.019
F_3	0.000	0.004	0.012	0.041	0.000	0.001	0.002	0.029	0.000	0.000	0.000	0.023
F_4	19.895	19.895	0.001	0.014	19.895	19.895	0.000	0.011	19.895	19.815	0.116	0.013
F_5	-1.031	-1.021	0.008	0.034	-1.031	-1.031	0.000	0.024	-1.031	-1.031	0.000	0.019
F_6	1.000	1.000	0.000	0.018	1.000	1.000	0.000	0.012	1.000	1.000	0.000	0.014
F_7	0.008	0.035	0.071	0.038	0.000	0.001	0.003	0.022	0.000	0.004	0.002	0.020
F_8	0.013	0.051	0.152	0.037	0.000	0.002	0.032	0.024	0.000	0.004	0.008	0.021

۴- الگوریتم ژنتیک کوانتومی برای مسئله PSP

$$(14)$$

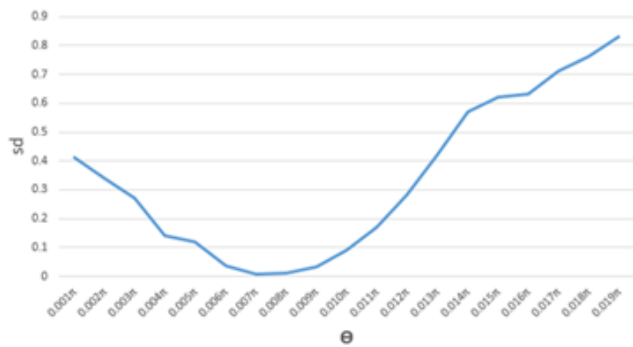
اجرای الگوریتم ژنتیک کوانتومی برای مسئله PSP برای ساختار دوبعدی و سه‌بعدی، بر اساس سیستم کوانتومی با استفاده از زبان برنامه‌نویسی C# انجام شده است. شبیه‌سازی‌ها بر روی سیستمی با مشخصات Intel Core i5، ۶ گیگابایت رم و ویندوز ۱۰ انجام شده است. ما الگوریتم ژنتیک کوانتومی پیشنهادی خود را با الگوریتم ژنتیک مرجع شماره [۱۱] برای ساختار دوبعدی و الگوریتم‌های ژنتیک مرجع شماره [۱۰، ۱۳] برای ساختار سه‌بعدی مقایسه کردیم.

جدول (۸)، پنج نمونه انتخاب شده دنباله آمینواسید HP را نشان می‌دهد که معیار استاندارد برای آزمایش توانایی الگوریتم‌ها برای یافتن جواب بهینه مسئله پیش‌بینی ساختار پروتئینی، هستند.

جدول ۸- پنج نمونه انتخاب شده دنباله آمینواسید HP

شماره	طول	دنباله پروتئین
-------	-----	----------------

معیار تابع تناسب برای دومین دنباله جدول ۸ را نشان می‌دهد. شکل ۶ نمودار انحراف $(0,001\pi \leq \theta \leq 0,019\mu)$ با مراحل 0.001 تغییر می‌کند. شکل ۶ نمودار انحراف



شکل ۶- انحراف معیار تابع تناسب برای دومین دنباله جدول ۸.

برای شبیه‌سازی الگوریتم ژنتیک کوانتومی ارائه شده برای مسئله PSP در ساختار دوبعدی، ما از نمونه دنباله آمینواسید HP جدول (۸) استفاده کرده‌ایم. اندازه جمعیت در این شبیه‌سازی برابر ۵۰ قرار داده شد. و از یک مقدار $(\pi 0,001)$ برای سه زاویه موجود در جدول جستجوی برای بروز رسانی ژن‌های کیوتريت استفاده شد. در جدول (۹) نتایج شبیه‌سازی و مقایسه نتایج با مراجع ۱۱ آمده است.

جدول ۹- نتایج شبیه‌سازی الگوریتم ژنتیک کوانتومی برای مسئله PSP در ساختار دوبعدی برای پنج دنباله انتخاب شده.

دنباله	GA	QGA
۱ زمان (ثانیه) چیدمان ایجاد شده مقدار جواب بهینه	- 30, 492 -9	19.069473505020142 1,117 -9
چیدمان	-	DLLURULUURDRURDDL
۲ زمان (ثانیه) چیدمان ایجاد شده مقدار جواب بهینه	- 30, 491 -9	39.961580753326416 2,192 -9
چیدمان	-	LLURURRRDRDLDDLULURR
۳ زمان (ثانیه) چیدمان ایجاد شده مقدار جواب بهینه	- 20, 400 -8	174.38035345077515 9,206 -8
چیدمان	-	UURULUULDDLULDDRDRUU
۴ زمان (ثانیه) چیدمان ایجاد شده مقدار جواب بهینه	- 301, 33, 9 -14	668.992345180511486 32,313 -14
چیدمان	-	ULDLULLLLDRURDRURDRDLDDL ULDLUL
۵ زمان (ثانیه) چیدمان ایجاد شده	- 126 .54 7	2,568.93060549316410624 114,081

مقدار جواب	-23	-23
بهینه		
چیدمان	-	DLDLULDLULLLURURDRURRULLLUL URRDRURDRURDRDLDDL

برای شبیه‌سازی الگوریتم ژنتیک کوانتومی ارائه شده برای مسئله PSP در ساختار سه‌بعدی، از نمونه دنباله آمینواسید HP جدول (۷) استفاده شد. اندازه جمعیت در این شبیه‌سازی برابر ۵۰ قرار داده شد. و از یک مقدار $(\pi 0,001)$ برای سه زاویه موجود در جدول جستجوی برای بروز رسانی ژن‌های کیوتريت استفاده شد. شرط پایان برای این شبیه‌سازی از نوع پایان بعد از تعداد تکرار معین قرار داده شد. در جدول (۱۰) نتایج شبیه‌سازی و مقایسه نتایج با مراجع ۱۰ آمده است.

جدول ۱۰- نتایج شبیه‌سازی الگوریتم ژنتیک کوانتومی برای مسئله PSP در ساختار سه‌بعدی برای پنج دنباله انتخاب شده.

دنباله	GA	دنباله پروتئین
۱ مقدار جواب بهینه چیدمان	-15 -	-11 DFDRURBLBUFRFLFUBLD
۲ مقدار جواب بهینه چیدمان	-13 -	-13 DRDLDBULFULBRBRFRULBLFF
۳ مقدار جواب بهینه چیدمان	-12 -	-9 RURULUUFFDBDFULBBBRULFF

۵- نتیجه‌گیری و کارهای آینده

در این پژوهش، ابتدا یک جدول جستجوی جدید برای بروز رسانی ژن‌های کیوتريت الگوریتم ژنتیک کوانتومی معرفی و مقایسه شد. در جدول جستجوی پیشنهادی، از گیت‌های چرخشی x ، y و z برای تغییر دامنه‌های ژن‌های کیوتريت استفاده شد که این امکان را می‌دهد که در مقایسه با پارامتر μ ، از ابزارهای متنوع‌تری برای تغییر دامنه ژن‌ها استفاده شد. طبق شبیه‌سازی انجام شده و نتایج به دست آمده، جدول جستجوی جدید برای تغییر دامنه‌های ژن‌های کیوتريت در مقایسه با پارامتر μ بهتر عمل می‌کند و می‌تواند به جواب بهینه سراسری سریع‌تر و دقیق‌تر دست پیدا کند. پس می‌توان نتیجه‌گیری کرد که استفاده از جدول جستجو برای بروز رسانی ژن‌های کیوتريت در الگوریتم ژنتیک کوانتومی بسیار کارآمد و بهتر از پارامتر μ است و می‌تواند جواب بهینه سراسری را سریع‌تر و دقیق‌تر پیدا کند.

در این پژوهش الگوریتم جدیدی را برای یافتن جواب بهینه سراسری چیدمان پروتئین HP با استفاده از الگوریتم ژنتیک کوانتومی پیشنهاد شد. در این پژوهش مسئله PSP در ساختار دوبعدی و سه‌بعدی، الگوریتم ژنتیک با تمام گام‌ها و مراحل آن، محاسبات کوانتومی و همچنین الگوریتم ژنتیک کوانتومی با تمام گام‌ها و مراحل آن، به‌طور کامل توضیح داده شد. گام‌ها و مراحل الگوریتم ژنتیک کوانتومی پیشنهادی برای مسئله PSP در ساختار دوبعدی و سه‌بعدی توضیح داده شد. همچنین برای بررسی و آزمایش عملکرد و کارایی الگوریتم ژنتیک کوانتومی، از پنج نمونه دنباله آمینواسید HP برای مسئله PSP در ساختار دوبعدی و سه‌بعدی استفاده شد.

با توجه به نتایج به دست آمده می‌توان گفت که الگوریتم ژنتیک کوانتومی پیشنهاد شده برای مسئله PSP در ساختار دوبعدی می‌تواند جواب بهینه سراسری را با تولید تعداد کروموزوم کمتر به دست آورد، که با توجه به این موضوع می‌توان گفت الگوریتم ژنتیک کوانتومی سریع‌تر از الگوریتم‌های ژنتیک ارائه شده برای مسئله PSP در ساختار دوبعدی عمل می‌کند، که علت اصلی آن استفاده از محاسبات کوانتومی

- [18]. Dill, Ken A. "Theory for the folding and stability of globular proteins." *Biochemistry* 24.6 (1985): 1501-1509.
- [19]. Luque-Chang, Alberto, et al. "首页 « 文章 » 文章详细信息." *Mathematical Problems in Engineering* 2018 (2018).
- [20]. Y. Sun, H. Xiong, Function optimization based on quantum genetic algorithm, *Research Journal of Applied Sciences, Engineering & Technology* 7 (1) (2014) 144-149.
- [21]. V. Tkachuk, Quantum genetic algorithm based on qutrits and its application, *Mathematical Problems in Engineering* 2018.

برای انجام محاسبات لازم مسئله است. اما برای مسئله PSP در ساختار سه‌بعدی به دلیل اینکه مراجع مربوط زمان و تعداد کروموزوم تولید برای رسیدن به جواب بهینه سراسری در مسئله PSP در ساختار سه‌بعدی را بیان نکرده‌اند، نمی‌توان این حرف زد و تنها می‌توان گفت الگوریتم ژنتیک کوانتومی می‌تواند جواب بهینه سراسری پنج نمونه دنباله آمینواسید HP را پیدا کند. در نتیجه پیشنهاد می‌شود که از الگوریتم ژنتیک کوانتومی، به دلیل سریع‌تر بودن آن در یافتن جواب بهینه سراسری مسئله، در مسئله PSP در ساختار دوبعدی استفاده شود.

۷- معرفی نویسندگان:



نام و نام خانوادگی: نیره مجد
 مرتبه علمی استادیار
 آدرس محل کار:
 دانشکده علوم مهندسی- پردیس دانشکدگان فنی - دانشگاه تهران - تهران- ایران
 تلفن: ۰۲۱۶۱۱۲۱۵۴
 پست الکترونیک naymajd@ut.ac.ir
 تحصیلات

- دکتری، ۱۳۸۵، فیزیک، دانشگاه صنعتی شریف
- کارشناسی ارشد، ۱۳۸۰، ذرات بنیادی، الزهرا
- کارشناسی، ۱۳۷۷، فیزیک کاربردی، الزهرا



نام و نام خانوادگی: معین عطاری
 دانشجوی فارغ‌التحصیل الگوریتم و محاسبات - مقطع کارشناسی ارشد
 آدرس محل کار:
 دانشکده علوم مهندسی- پردیس دانشکدگان فنی - دانشگاه تهران - تهران- ایران
 تلفن: ۰۲۱۶۱۱۲۱۵۴
 پست الکترونیک @ut.ac.ir
 تحصیلات

- کارشناسی ارشد، ۱۴۰۱، الگوریتم و محاسبات، دانشگاه تهران
- کارشناسی، ۱۳۹۸، مهندسی کامپیوتر، دانشگاه سجاد مشهد

۶- مآخذ

- [1]. Davis, Ian W., and David Baker. "RosettaLigand docking with full ligand and receptor flexibility." *Journal of molecular biology* 385.2 (2009): 381-392.
- [2]. Raman, Srivatsan, et al. "Accurate automated protein NMR structure determination using unassigned NOESY data." *Journal of the American Chemical Society* 132.1 (2010): 202-207.
- [3]. Dobson, Christopher M. "Principles of protein folding, misfolding and aggregation." *Seminars in cell & developmental biology*. Vol. 15. No. 1. Academic Press, 2004.
- [4]. Dobson, Christopher M. "Protein folding and misfolding." *Nature* 426.6968 (2003): 884-890.
- [5]. Anfinsen, Christian B. "Principles that govern the folding of protein chains." *Science* 181.4096 (1973): 223-230.
- [6]. James, Leo C., and Dan S. Tawfik. "Conformational diversity and protein evolution—a 60-year-old hypothesis revisited." *Trends in biochemical sciences* 28.7 (2003): 361-368.
- [7]. Crescenzi, Pierluigi, et al. "On the complexity of protein folding." *Journal of computational biology* 5.3 (1998): 423-465.
- [8]. Wetlaufer, Donald B. "Nucleation, rapid folding, and globular intrachain regions in proteins." *Proceedings of the National Academy of Sciences* 70.3 (1973): 697-701.
- [9]. Bui, Thang N., and Gnanasekaran Sundarraj. "An efficient genetic algorithm for predicting protein tertiary structures in the 2D HP model." *Proceedings of the 7th annual conference on Genetic and evolutionary computation*. 2005.
- [10]. Khimasia, Mehul M., and Peter V. Coveney. "Protein structure prediction as a hard optimization problem: the genetic algorithm approach." *Molecular Simulation* 19.4 (1997): 205-226.
- [11]. Custódio, Fábio Lima, Helio JC Barbosa, and Laurent Emmanuel Dardenne. "A multiple minima genetic algorithm for protein structure prediction." *Applied Soft Computing* 15 (2014): 88-99.
- [12]. Sadek, Bouroubi, and Nabil Boumedine. "A new hybrid genetic algorithm for protein structure prediction on the 2Dtriangular lattice." *Turkish Journal of Electrical Engineering and Computer Sciences* 29.2 (2021): 499-513.
- [13]. Wang, Ting, and Xiaolong Zhang. "3D Protein structure prediction with genetic tabu search algorithm in Off-Lattice AB model." *2009 Second International Symposium on Knowledge Acquisition and Modeling*. Vol. 1. IEEE, 2009.
- [14]. Atari, Moein, and Nayereh Majd. "2D HP protein folding using quantum genetic algorithm." *2022 27th International Computer Conference, Computer Society of Iran (CSICC)*. IEEE, 2022.
- [15]. Bryant, Patrick. "Deep learning for protein complex structure prediction." *Current opinion in structural biology* 79 (2023): 102529.
- [16]. Stahl, Kolja, et al. "Protein structure prediction with in-cell photo-crosslinking mass spectrometry and deep learning." *Nature Biotechnology* 41.12 (2023): 1810-1819.
- [17]. Fang, Xiaomin, et al. "A method for multiple-sequence-alignment-free protein structure prediction using a protein language model." *Nature Machine Intelligence* 5.10 (2023): 1087-1096.

3. James and Twafik

4. Crescenzi

5. Wetlaufer

6. Bui

7. Khimasia

8. Fábio

9. Multiple minima GA

10. Zhang

11. Genetic tabu search algorithm

12. Lattice proteins

13. Bead

14. Look-up table

15. Co-look-up table