نشریه علمی «نوآوری‌های فناوری اطلاعات و ارتباطات کاربردی»
سال اول، شماره ۳، پاییز ۱۴۰۰؛ ص ۱۲- ۱

علمی – تخصصی

الگوریتم تکاملی مبتنی بر مدل با استفاده از خوشه‌بندی فازی C– میانگین و تحلیل مؤلفه‌های اصلی

پژمان غلام‌نژاد*

دکتری تخصصی کامپیوتر هوش مصنوعی- دانشکده مهندسی کامپیوتر و فناوری اطلاعات-دانشگاه علوم و فنون هوایی شهید ستاری- تهران- ایران
(دریافت: ۱۴۰۰/۰۴/۱۶، پذیرش: ۱۴۰۰/۱۰/۲۱)

چکیده

ساختار عملگرها در بیشتر الگوریتم‌های تکاملی چند هدفه سنتی بر اساس قوانین ثابت اکتشافی مانند تقاطع و جهش است که قادر به یادگیری ساختارها یا خصوصیات مسائل بهینه‌سازی نیستند. برای تجهیز الگوریتم‌های تکاملی به توانایی‌های یادگیری، اخیراً الگوریتم‌های تکاملی مبتنی بر مدل ارائه‌شده است. در الگوریتم‌های تکاملی مبتنی بر مدل، عملگرهای ابتکاری با مدل‌های یادگیری ماشین مانند مدل‌های آموزش و نمونه جایگزین می‌شوند. در این مقاله یک الگوریتم تکاملی چند هدفه مبتنی بر مدل ارائه‌شده است که در آن هر نسل، یک ناحیه محتمل از فضای جستجو، به‌وسیله یک مدل احتمالی، می‌شود. بر روی ناحیه محتمل در فضای جستجو، که همان نقاط غالب، با مرتبه بهتر می‌باشند، خوشه‌بندی یا بر اساس روش‌های فازی انجام می‌گیرد و یا این‌که بر روی نقاط غالب با مرتبه اول، یک عملگر انتخاب مسابقه‌ای ازدحام، صورت می‌پذیرد تا نقاط بافاصله نزدیک به هم، حذف شوند و نتیجه حاصل به‌عنوان نقاط مرکزی خوشه‌ها در نظر گرفته می‌شود و سپس، خوشه‌بندی بر اساس نزدیک‌ترین همسایه‌ها، صورت می‌پذیرد. از الگوریتم تجزیه تحلیل مؤلفه‌های اصلی، که بهترین روش برای کاهش ابعاد داده به‌صورت خطی می‌باشد، برای مدل‌سازی، استفاده‌شده است. راه‌حل‌های جدید از مدل ساخته‌شده، بر اساس یک توزیع نرمال، به‌دست‌آمده می‌آیند. روش پیشنهادی مورد آزمایش قرارگرفته است و نتایج حاصل از آن با روش الگوریتم‌های ژنتیک مرتب‌سازی نامغلوب، مقایسه شده است. نتایج نشان می‌دهد که این روش سریع‌تر از روش‌های قبلی می‌باشد و با تعداد کمتری تکرار و ارزیابی توابع، نتایج بهتری به‌دست می‌آید.

کلیدواژه‌ها: عملگر انتخابی ازدحام، الگوریتم‌های تکاملی چند هدفه مبتنی بر مدل، خوشه‌بندی فازی ، بهینه‌سازی چند هدفه، تحلیل مولفه‌های اساسی .

۱- مقدمه

۱-۱- بیان مساله

طی سه دهه اخیر الگوریتم‌های تکاملی چند هدقه بسیاری ارائه و پیشنهادشده‌اند که می‌توان آن‌ها را در سه گروه دسته‌بندی نمود [۱]: روش‌های بر پایه وزن‌دهی مجتمع[1]، روش‌های برپایِ غلبگی[2]، و نگرش برپایِ کارایی شاخص[3]. روش‌های برپایِ وزن‌دهی مجتمع، به‌عنوان نگرش تجزیه شناخته می‌شوند که در آن یک مساله بهینه‌سازی چند هدفه به یک تعداد از مسائل بهینه‌سازی تک هدفه، با استفاده از یک تعداد ترکیبات وزنی، که به‌صورت تصادفی ایجاد می‌شوند، تجزیه می‌شود.

در الگوریتم‌های تکاملی چند هدفه بر پایه غلبگی، استفاده از نخبه‌گرایی به‌صورت قابل‌ملاحظه‌ای، همگرایی در این الگوریتم‌ها را افزایش می‌دهد. در این گروه از الگوریتم‌ها با افزایش تعداد اهداف، به‌صورت قابل‌ملاحظه‌ای همگرایی کاهش می‌یابد که عمدتاً به‌دلیل کاهش فشار انتخاب است.

یکی دیگر از ایده‌ها، اختصاص یک برازش به افراد یک الگوریتم بهینه‌سازی چند هدفه برپایِ شاخص کارایی است. در میان این الگوریتم‌ها، اَبَر حجم[4]، می‌تواند هم دقت و هم تنوع مجموعه راه‌حل‌های نامغلوب را محاسبه نماید. چالش اصلی این گروه از الگوریتم‌ها، پیچیدگی محاسباتی برای محاسبه شاخص کارایی است، مخصوصاً زمانی که تعداد اهداف زیاد است.

بیشتر الگوریتم‌های ذکرشده متمرکز بر بسط و توسعه یک برآورد برازش مؤثر و یا راهبرد انتخاب، منطبق بر یک الگوریتم

* رایانامه نویسنده مسئول: pezhman.gholamnezhad@gmail.com
[1] Weighted Aggregation based method
[2] Dominance based method
[3] Indicator based method
[4] Hypervolume

تکاملی تک هدفه، برای حل مسائل بهینه‌سازی چند هدفه تکاملی است. توجه بسیار زیادی به طراحی راهبرد بازتولید مؤثر که به‌صورت واضح و منظم و ارتباط را در توزیع راه‌حل‌های بهینه پرتو، بپردازد، نشده است. اگرچه بسیاری از راهبردهای محلی، روش‌های وزن‌دهی مجتمع پویا، راهبرد جفت‌گیری محدود[۱]، به‌صورت ضمنی از خصوصیات مرتبط، برای افزایش مؤثر جستجو، بهره برده‌اند، اما عملگرها (تقاطع[۲] و جهش[۳]) در اکثر الگوریتم‌های تکاملی سنتی، بر اساس قوانین اکتشافی یا راهبردهای ثابت ساخته شده‌اند که قادر به یادگیری ساختارها یا خصوصیات مسائل بهینه‌سازی چند هدفه نمی‌باشند.

به‌منظور تجهیز الگوریتم‌های تکاملی به توانایی‌های یادگیری اخیراً الگوریتم‌های تکاملی مبتنی بر مدل[۴] در سه گروه ارائه‌شده است[۲]:

۱- الگوریتم‌های تخمین توزیع[۵] که مبتنی بر تخمین توزیع راه‌حل‌های کاندید می‌باشد.

۲- الگوریتم‌های مدل‌سازی معکوس[۶] که بر پایه ایجاد مدل‌های معکوس از فضای هدف به فضای تصمیم می‌باشد.

۳- الگوریتم‌های مدل‌سازی جایگزین[۷] که هنگام حل مسائل دارای بار محاسباتی سنگین استفاده می‌شوند. الگوریتم‌های تخمین توزیع به سه گروه زیر دسته‌بندی می‌شوند: الگوریتم‌های تخمین توزیع تک متغیره[۸]. الگوریتم‌های تخمین توزیع چند متغیره[۹]. الگوریتم‌های تخمین توزیع چند هدفه[۱۰].

برای تخمین توزیع راه‌حل‌های کاندید در فضای تصمیم، همبستگی متغیر عاملی اساسی است که باید در مدل‌سازی موردتوجه قرار گیرد. یک روش ساده اتخاذ مدل‌های تک متغیره است که بر پایه مستقل بودن متغیرهای تصمیم است. بر این اساس توزیع احتمال یک راه حل کاندیدِ $x=(x_1,x_2,\ldots,x_n)$ را می‌توان به‌صورت زیر تجزیه نمود:

$$p(x)=p(x_1)p(x_2)\ldots p(x_n) \qquad (۱)$$

که در آن $p(x)$ توزیع احتمال محلول کاندید x و $p(x)$ توزیع احتمال متغیر تصمیم x_i است. الگوریتم‌های تخمین توزیع‌هایی که از چنین مدل‌های توزیع یک متغیره استفاده می‌کنند، به‌عنوان الگوریتم تخمین توزیع تک متغیره شناخته می‌شوند.

الگوریتم‌های تخمین توزیع چند متغیره مبتنی بر در نظر گرفتن اثر متقابل بین زوج متغیرهای تصمیم می‌باشد. به‌طورکلی توزیع احتمال یک راه حل کاندیدِ $x=(x_1,x_2,\ldots,x_n)$ را می‌توان بر اساس مدل احتمال شرطی به‌صورت زیر بیان نمود:

$$p(x)=p(x_{i_1}|x_{i_2})p(x_{i_2}|x_{i_3})\ldots p(x_{i_{n-1}}|x_{i_n})p(x_{i_n}) \qquad (۲)$$

که در آن $p(x)$ توزیع احتمال راه‌حل کاندید x و $p(x_{i_j}|x_{i_{j+1}})$ توزیع احتمال شرطی x_{i_j} با توجه به $x_{i_{j+1}}$ است.

الگوریتم‌های تخمین توزیع چند هدفه به جای دستیابی به یک بهینه عمومی، مجموعه‌ای از راه‌حل‌های بهینه را به‌عنوان تقریب با جبهه پارتو[۱۱] (و همچنین مجموعه پارتو[۱۲]) به‌دست آورند.

۱-۲- اهمیت و ضرورت تحقیق

الگوریتم‌های تخمین توزیع چند هدفه به‌عنوان متداول‌ترین الگوریتم‌های تکاملی مبتنی بر مدل، در دهه گذشته پیشرفت‌های چشم‌گیری داشته‌اند. مزیت اصلی الگوریتم‌های تخمین توزیع، توانایی بالقوه سازگاری با محیط برازش[۱۳] و یادگیری ساختار مساله می‌باشد.

مهم‌ترین موضوعی که در الگوریتم‌های تکاملی باعث کند شدن روند یافتن پاسخ بهینه می‌شود، تکرارهای متوالی و به‌نوعی حلقهٔ اصلی تولید نسل است. علاوه بر این هر تکرار حلقه بعضاً دارای سربار زمانی زیادی است (ناشی از عملگرهای جهش، تقاطع و برخی اوقات مرتب کردن کروموزوم‌ها). ارزش این نوع الگوریتم‌ها هم در این است که در بسیاری از موارد از این عملگرها استفاده نمی‌کند به عبارت بهتر، مفهوم پایهٔ الگوریتم‌های تخمین توزیع

[1] Hypervolume
[2] Crossover
[3] Mutation
[4] Model-Based Evolutionary Algorithms
[5] Estimation of distribution
[6] Inverse Modeling
[7] Surrogate Modeling
[8] Univariate EDAs
[9] Multivariate EDAs
[10] Multi-Objectve EDA
[11] Pareto Front (PF)
[12] Pareto Set (PS)
[13] fitness

درواقع در مبنای تکامل طبیعی و عملگرهای جهش و تقاطع بنانهاده نشده است.

این الگوریتم‌ها برای مسائلی مفید است که دارای برخی خصوصیات خاص می‌باشند. این الگوریتم‌ها توزیع راه‌حل‌های کاندید امیدوار را با آموزش و نمونه‌برداری از مدل‌ها در فضای تصمیم تخمین می‌زنند.

۱-۳- هدف تحقیق

در این پژوهش یک الگوریتم بهینه‌سازی چند هدفه تکاملی مبتنی بر مدل (الگوریتم‌های تخمین توزیع چند هدفه) با استفاده از خوشه‌بندی فازی و تحلیل مؤلفه اصلی ارائه می‌شود که در هر نسل، یک ناحیه محتمل از فضای جستجو، به‌وسیله یک مدل احتمالی، مدل‌سازی می‌شود. بر روی ناحیه محتمل در فضای جستجو، که همان نقاط غالب، با مرتبه بهتر می‌باشند، خوشه‌بندی به دو صورت انجام می‌شود (یا بر اساس روش‌های فازی انجام می‌گیرد و یا این‌که بر روی نقاط غالب با مرتبه اول، یک عملگر انتخاب مسابقه‌ای ازدحام، صورت می‌پذیرد تا نقاط با فاصله نزدیک به هم، حذف شوند و نتیجه حاصل به‌عنوان نقاط مرکزی خوشه‌ها در نظر گرفته می‌شود و سپس، خوشه‌بندی بر اساس نزدیک‌ترین همسایه‌ها، صورت می‌پذیرد.) از الگوریتم تجزیه تحلیل مؤلفه‌های اصلی، که بهترین روش برای کاهش ابعاد داده به‌صورت خطی می‌باشد، برای مدل‌سازی، استفاده‌شده است. راه‌حل‌های جدید از مدل ساخته شده، بر اساس یک توزیع نرمال، به‌دست می‌آیند.

۱-۴- پیشینه تحقیق

در الگوریتم بهینه‌سازی چند هدفه بیزی[1]، عملگر انتخاب مبتنی بر یک ε -بایگانی[2] است، که یک مجموعه‌ی حداقلی از راه‌حل‌های کاندید که بر تمامی راه‌حل‌ها را غالب است، در تمام نسل‌ها نگهداری می‌شود [۳]. در الگوریتم تکاملی تخمین چگالی تکراری چند هدفه مبتنی بر مخلوط ساده[3]، فشار انتخاب دوفازی اعمال می‌شود، که فشار انتخاب توسط یک پارامتر δ تنظیم می‌شود[۴]. الگوریتم چند هدفه بیزی سلسله مراتبی[4]، عملگر انتخاب در الگوریتم ژنتیک رتبه‌بندی نامغلوب[5] را با روش خوشه‌بندی k-میانگین[6] ترکیب می‌کند [۵]. برخلاف بیشتر الگوریتم‌های تخمین توزیع چند هدفه که عملگرهای انتخاب جدیدی را ارائه می‌دهند، تخمین چند هدفه الگوریتم توزیع مبتنی بر مدل منظم[7]، یک عملگر تولیدمثل جدید را ایجاد می‌کند. این روش با استفاده از روش محلی تحلیل مؤلفه‌های اساسی، ابعاد بردارهای تصمیم‌گیری را کاهش می‌دهد و سپس از راه‌حل‌های جدید کاندید در فضای نهفته نمونه‌برداری می‌کند [۶]. الگوریتم تخمین چگالی چند هدفه بر پایه مدل منظم یک عملگر بازترکیبی را با مدل‌سازی منظم در توزیعِ راه‌حل‌های پرتو بهینه، جایگزین می‌کند. در این الگوریتم، در هر نسل، یک ناحیه محتمل در فضای تصمیم، با یک توزیع احتمالی مدل می‌شود و از الگوریتم تحلیل مولفه اساسی محلی[8] برای ساختن مدل استفاده می‌کند و راه‌حل‌های جدید از مدل ساخته‌شده، نمونه‌برداری می‌شود و از یک انتخاب بر پایه مرتب‌سازی نامغلوب، به‌منظور انتخاب راه‌حل‌ها برای نسل بعدی، استفاده می‌شود. درواقع، در هر نسل، فضای تصمیم به خوشه، تقسیم‌بندی می‌شود و از الگوریتمِ بعدی تحلیل مولفه اساسی محلی استفاده می‌کند و با استفاده از بردار ویژه[9] و بزرگ‌ترین مقدار ویژه[10] به‌دست‌آمده، برای مدل‌سازی در فضای تصمیم، استفاده می‌شود [۷]. در الگوریتم ژنتیک مرتب‌سازی نامغلوب [۸]، ابتدا جمعیت فرزندان، با استفاده از جمعیت والدین ساخته می‌شود و به‌جای پیدا کردن جواب‌های نامغلوب از فرزندان، ابتدا دو جمعیت والدین و فرزندان را با یکدیگر ترکیب کرده و سپس از یک مرتب‌سازی نامغلوب، برای دسته‌بندی تمام جمعیت استفاده می‌شود. در الگوریتم تکاملی چند هدفه بر پایه مدل‌سازی معکوس [۹]، توزیعی از والدین فعلی در فضای هدف و فضای تصمیم وجود دارد. توزیع‌های جمعیت فعلی، در فضای تصمیم با مشخص می‌شود و در فضای هدف با مشخص می‌گردد و توزیع احتمال شرطی ، می‌تواند به‌وسیله یک مدل احتمالی معکوس، تقریب زده شود که راه‌حل‌های نامغلوب را از فضای هدف، یعنی پرتو نما[11] که با Yp نمایش داده می‌شود، را به فضای تصمیم یعنی مجموعه پرتو[12]، که با Xp نمایش داده می‌شود، نگاشت می‌کند. برای ایجاد

[1] Bayesian multi-objective optimization algorithm (BMOA)
[2] archive
[3] naive mixture-based multi-objective iterated density estimation evolutionary algorithm (MIDEA)
[4] multi-objective hierarchical BOA(mohBOA)

[5] NSGA-II
[6] K-means
[7] regularity model-based multi-objective estimation of distribution algorithm (RM-MEDA)
[8] Local Principal Component Analysis
[9] Eigen vector
[10] Largest eigenvalue
[11] Pareto Front
[12] Pareto Set

فرزندان، این الگوریتم با ایجاد نمونه‌ها در فضای هدف که با Yo، مشخص می‌شوند، بر پایه اطلاعات مقادیر هدف، در جمعیت والدین فعلی شروع می‌کند. هنگامی‌که مجموعه فرزندان Yo، ایجاد شدند، به‌صورت برگشتی با ، با استفاده از تئوری بیز، به فضای تصمیم نگاشت می‌شوند.

۲- مفاهیم، نظریه‌های مربوطه و فرضیه‌های تحقیق

الگوریتم‌های تکاملی به‌عنوان ابزاری مؤثر برای کاوش در جبهه‌های بهینه پارتوِ همگرا[۱] و متنوع[۲] در مسائل بهینه‌سازی چند هدفه ارائه‌شده‌اند که به‌طورکلی شامل دو یا سه هدف متناقض است. مسائل بهینه‌سازی چند هدفه، مربوط به حل M اهدافِ متضاد به‌طور هم‌زمان در مسائلی است که در آن مقدار M، برابر با دو یا سه است[۱۰]. به‌طورکلی، یک مساله بهینه‌سازی چند هدفه فرمول زیر را دارد که توسط معادله (۱) توصیف‌شده است.

$$\begin{cases} f(x) = [f_1(x), \ldots, f_M(x)] \\ s.t. \quad x \in X \end{cases} \qquad (۳)$$

که $X \in \Re^n$، فضای تصمیم[۳] است و $f : X \to \Re^M$ ، فضای هدف[۴] است و فرض بر این است که f، یک مساله کمینه‌سازی است.

۲-۱- خوشه‌بندی فازی c-میانگین[۵]

در مسائل شناسایی الگو منظور از خوشه، مجموعه‌ای از داده‌ها است که به علت شباهت زیادی که به هم دارند در یک دسته قرارگرفته‌اند. در خوشه‌بندی سعی می‌شود جمعی از داده‌ها به‌صورت بدون ناظر به خوشه‌هایی تقسیم شوند که شباهت داده‌های درون هر خوشه حداکثر و شباهت بین داده‌های درون خوشه‌های مختلف حداقل گردد. یکی از رایج‌ترین روش‌های خوشه‌بندی، خوشه‌بندی فازی c-میانگین است [۱۱].

این الگوریتم از یک شیوه ساده برای خوشه‌بندی کردن یک مجموعه داده در یک تعداد خوشه از پیش مشخص‌شده، استفاده می‌کند. ایده اصلی تعریف k مرکز برای هر یک از خوشه‌ها می‌باشد. این مراکز بایستی با دقت زیاد انتخاب شوند، زیرا مراکز مختلف، نتایج مختلف را به‌وجود می‌آورند.

بنابراین بهترین انتخاب قرار دادن آن‌ها (مراکز) در فاصله هر چه بیشتر از یکدیگر می‌باشد . قدم بعدی تخصیص هر الگو به نزدیک‌ترین مرکز می‌باشد. وقتی همه‌ی نقاط به مراکز موجود تخصیص داده شدند، مرحله اول تکمیل‌شده است و یک گروه‌بندی اولیه انجام‌شده است. در این مرحله نیاز داریم که k مرکز جدید برای خوشه‌های مرحله قبل محاسبه کنیم. بعد از تعیین k مرکز جدید، مجدداً داده‌ها را به مراکز مناسب تخصیص می‌دهیم. این مراحل را آن‌قدر تکرار می‌کنیم که دیگر k مرکز، جابه‌جا نشوند.

۲-۲- خوشه‌بندی گاستافسون-کِسِل[۶]

این الگوریتم بسط یافته الگوریتم c-میانگین فازی می‌باشد [۱۲] که در آن، برای پیدا کردن خوشه‌هایی با اشکال هندسی متفاوت، بر روی مجموعه داده‌ها، فاصله نرمال تطبیقی[۷]، به‌کار گرفته‌شده است [۱۳]. الگوریتم میانگین فازی، خوشه‌ها را به‌صورت کرَوی جستجو می‌کند، اما این روش خوشه‌ها را به‌صورت بیضی‌شکل پیدا می‌کند. در این روش هر خوشه به‌وسیله مرکز و ماتریس کوواریانس، مشخص می‌شود. تابع هدف در این الگوریتم به‌صورت رابطه زیر، تعریف می‌شود:

$$J(x;U,V,A) = \sum_{i=1}^{c}\sum_{k=1}^{N} (\mu_{ik})^m D^2_{ikA_i} \qquad (۴)$$

که در آن:

$$D^2_{ikA_i} = (x_k - v_i)^T A_i (x_k - v_i) \qquad (۵)$$

می‌باشد. این الگوریتم از فاصله نرمال ماهال[۸]، استفاده می‌کند. در این دو رابطه، μ_{ik} درجه عضویت هر فرد به هر خوشه می‌باشد. $D^2_{ikA_i}$ فاصله نرمال ماهال بین نقطه و خوشه k ، می‌باشد. A ماتریس کوواریانس خوشه‌ها می‌باشد. V مراکز خوشه‌ها است.

[1] Pareto optimal fronts
[2] diversified
[3] Decision Space
[4] Objective space
[5] Fuzzy C-mean Clustering (FCM)
[6] Gustafson - Kessel
[7] adaptive distance norm
[8] Mahalanibis distance norm

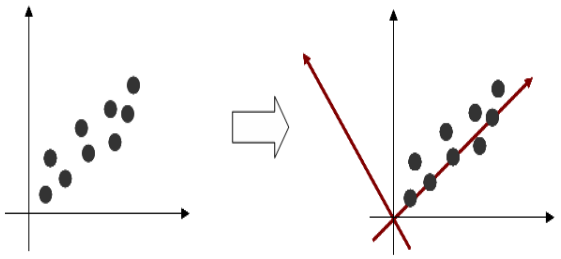
۲-۳- تحلیل مؤلفه اساسی[1]

بسترهای داده‌ای که دارای ابعاد زیادی هستند علیرغم فرصت‌هایی که به‌وجود می‌آورند، چالش‌های محاسباتی زیادی را ایجاد می‌کنند. یکی از مشکلات داده‌های با ابعاد زیاد این است که در بیشتر مواقع تمام ویژگی‌های داده‌ها برای یافتن دانشی که در داده‌ها نهفته است مهم و حیاتی نیستند. به همین دلیل در بسیاری از زمینه‌ها کاهش ابعاد داده یکی از مباحث قابل‌توجه باقی‌مانده است. روش‌های کاهش ابعاد داده به دودسته تقسیم می‌شوند:

روش‌های مبتنی بر استخراج ویژگی[۱۴]: این روش‌ها یک فضای چندبعدی را به یک فضای با ابعاد کمتر نگاشت می‌کنند. درواقع با ترکیب مقادیر ویژگی‌های موجود، تعداد کمتری ویژگی به‌وجود می‌آورند به‌طوری‌که این ویژگی‌ها دارای تمام (یا بخش اعظمی از) اطلاعات موجود در ویژگی‌های اولیه باشند. این روش‌ها به دودسته‌ی خطی و غیرخطی تقسیم می‌شوند.

روش‌های مبتنی بر انتخاب ویژگی[۱۵]: این روش‌ها سعی می‌کنند با انتخاب زیرمجموعه‌ای از ویژگی‌های اولیه، ابعاد داده‌ها را کاهش دهند. در پاره‌ای از اوقات تحلیل‌های داده‌ای نظیر طبقه‌بندی بر روی فضای کاسته شده نسبت به فضای اصلی بهتر عمل می‌کند.

روش تحلیل مؤلفه اساسی [۱۰] بهترین روش برای کاهش ابعاد داده به‌صورت خطی می‌باشد. یعنی با حذف ضرایب کم‌اهمیت به‌دست‌آمده از این تبدیل، اطلاعات ازدست‌رفته نسبت به روش‌های دیگر کمتر است. البته کاربرد تحلیل مؤلفه اساسی محدود به کاهش ابعاد داده نمی‌شود و در زمینه‌های دیگری مانند شناسایی الگو و تشخیص چهره نیز مورداستفاده قرار می‌گیرد. در این روش محورهای مختصات جدیدی برای داده‌ها تعریف‌شده و داده‌ها بر اساس این محورهای مختصات جدید بیان می‌شوند. اولین محور باید در جهتی قرار گیرد که واریانس داده‌ها بیشترین شود (یعنی در جهتی که پراکندگی داده‌ها بیشتر است). دومین محور باید عمود بر محور اول به‌گونه‌ای قرار گیرد که واریانس داده‌ها بیشترین شود. به همین ترتیب محورهای بعدی عمود بر تمامی محورهای قبلی به‌گونه‌ای قرار می‌گیرند که داده‌ها در آن جهت دارای بیشترین پراکندگی باشند. در شکل (۱) این مطلب برای داده‌های دوبعدی نشان داده شده است.



شکل (۱): انتخاب محورهای جدید برای داده‌های دوبعدی

مراحل الگوریتم تحلیل مؤلفه‌های اساسی به شرح زیر می‌باشد:

۱- انتخاب داده. ۲- کم کردن میانگین از داده‌ها (میانگین هر بُعد را از مقادیر آن بُعد کم می‌کنیم تا میانگین داده‌ها در هر بُعد صفر شود). ۳- محاسبه ماتریس کوواریانس ۴- محاسبه بردارهای ویژه و مقدارهای ویژه. ۵- انتخاب مؤلفه‌ها و ساختن بردار ویژگی. در این مرحله مفهوم کاهش ابعاد داده انجام می‌شود. بردارهای ویژه‌ای که در مرحله‌ی قبل به‌دست آوردیم را بر اساس مقادیر ویژه‌ی آن‌ها از بزرگ به کوچک مرتب می‌کنیم (توجه داشته باشید که مقادیر ویژه‌ی ماتریس کوواریانس همگی بزرگ‌تر یا مساوی صفر هستند). بدین ترتیب مؤلفه‌های داده‌ها از پراهمیت به کم‌اهمیت مرتب می‌شوند. در اینجا اگر بخواهیم ابعاد داده‌ها را کاهش دهیم می‌توانیم مؤلفه‌های کم‌اهمیت را حذف نماییم. البته این کار با از دست دادن مقدار کمی اطلاعات همراه است.

۲-۴- الگوریتم ژنتیک مرتب‌سازی نامغلوب[2]

در الگوریتم ژنتیک مرتب‌سازی نامغلوب [۸]، ابتدا جمعیت فرزندان، با استفاده از جمعیت والدین ساخته می‌شود و به‌جای پیدا کردن جواب‌های نامغلوب از فرزندان، ابتدا دو جمعیت والدین و فرزندان را با یکدیگر ترکیب کرده و سپس از یک مرتب‌سازی نامغلوب، برای دسته‌بندی تمام جمعیت استفاده می‌شود.

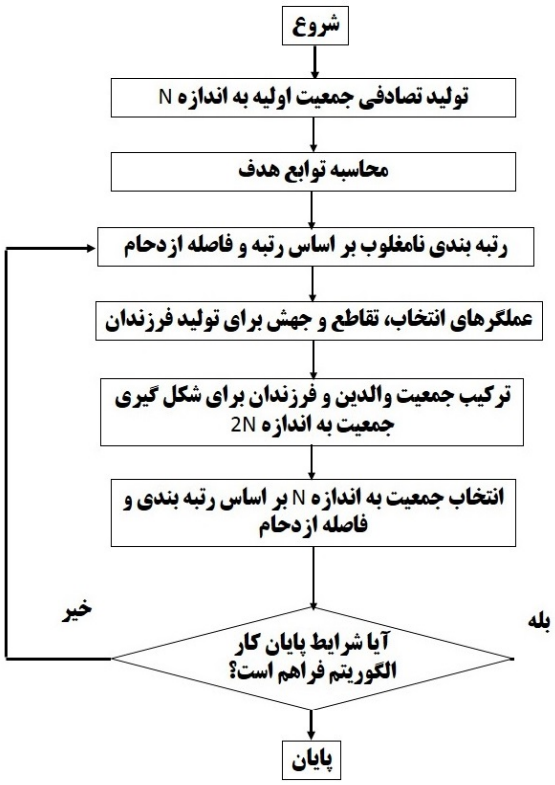
در این شیوه یک مقایسه عمومی بر روی فرزندان و والدین انجام می‌گیرد و پس از ایجاد صف‌های نامغلوب به ترتیب اولویت (اولویت صف‌ها نسبت به هم)، جمعیت بعدی یکی‌یکی از این صف‌ها پر می‌شود. پر کردن جمعیت جدید با بهترین صف نامغلوب شروع می‌شود و سپس به ترتیب با دومین صف نامغلوب و همین‌طور سومین و الی آخر، تا زمانی که جمعیت جدید پر شود، ادامه پیدا می‌کند. از آن‌جا که اندازه جمعیت ما دو برابر است، تمام اعضا در جمعیت جدید قرار نمی‌گیرند و بایستی

[1] Principal Component Analysis

[2] Non-dominated Sorting Genetic Algorithm (NSGA-II)

جمعیت باقیمانده را حذف کنیم. در مورد جواب‌هایی که در صف آخر با استفاده از عملگر نخبه‌گرایی از بین می‌رود، باید مهارت بیشتری به‌کار برده و جواب‌هایی را که در ناحیه ازدحام کمتری قرار دارند، حفظ کرد. درواقع برای رعایت اصل چگالی در بین جواب‌ها، جواب‌هایی که در ناحیه ازدحام کوچک‌تری هستند، برای پر کردن جمعیت جدید در اولویت قرار دارند.

چارچوب کلی این الگوریتم در شکل (۲) بیان‌شده است:



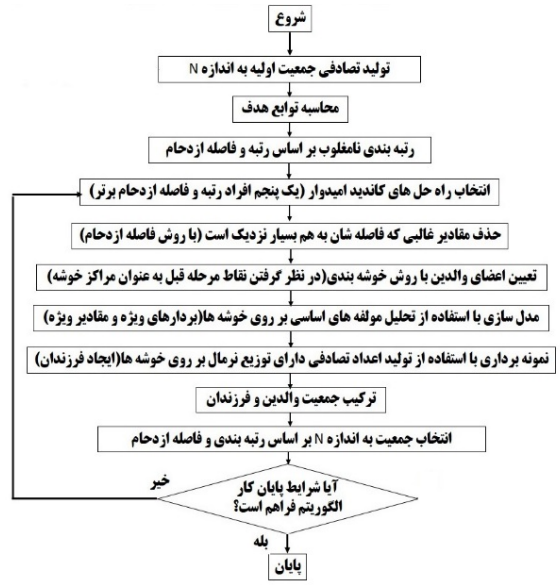
شکل (۲): چارچوب کلی الگوریتم ژنتیک مرتب‌سازی نامغلوب

۳- روش پیشنهادی

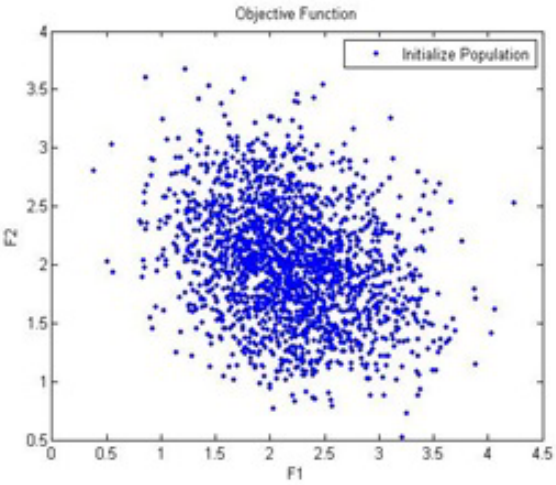
در این پژوهش یک الگوریتم بهینه‌سازی چند هدفه تکاملی مبتنی بر مدل (الگوریتم‌های تخمین توزیع چند هدفه) با استفاده از خوشه‌بندی فازی و تحلیل مؤلفه اصلی ارائه می‌شود.

چارچوب کلی الگوریتم پیشنهادی در شکل (۳) بیان‌شده است.

جمعیت اولیه شامل مقادیر n -بعدی می‌باشد و هر یک از افراد با x نمایش داده می‌شوند که به‌صورت تصادفی در فضای تصمیم تولیدشده‌اند. سپس مقادیر تابع برازش برای جمعیت اولیه محاسبه می‌گردد. شکل (۴) این امر را برای یک تابع برازش دوبعدی نمایش می‌دهد.



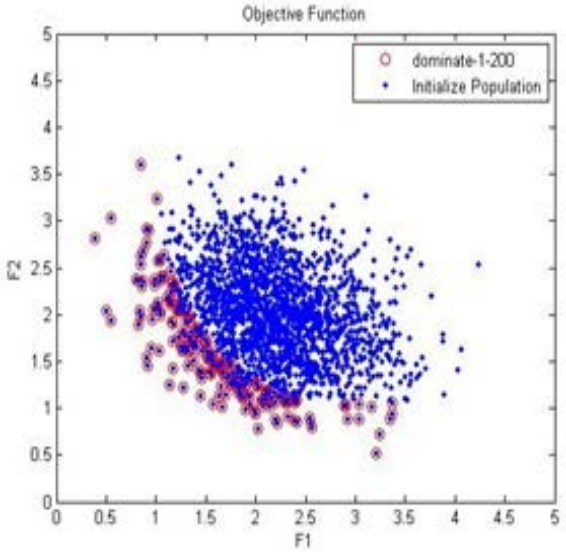
شکل (۳): چارچوب کلی الگوریتم پیشنهادی



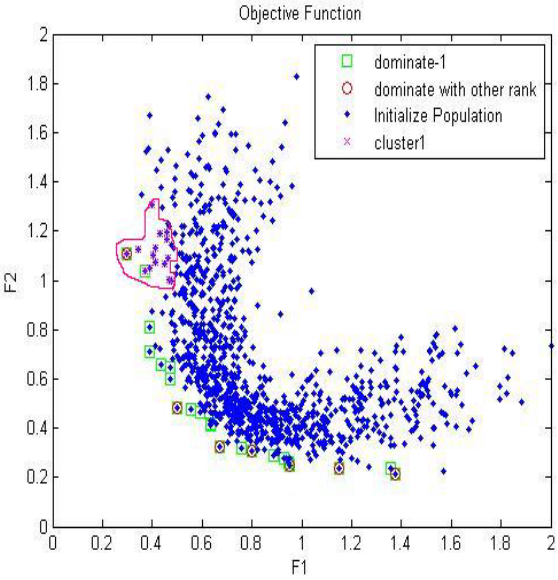
شکل (۴): مقادیر تابع برازش دوبعدی برای جمعیت اولیه

سپس رتبه‌بندی افراد نامغلوب بر اساس رتبه و فاصله ازدحام بر روی مقادیر تابع برازش صورت می‌پذیرد و افراد به‌عنوان راه-حل‌های کاندید امیدوار انتخاب می‌شوند که این امر در شکل (۵) نشان داده‌شده است.

سپس غالب‌های نزدیک به یکدیگر را حذف نموده و اگر فاصله غالب‌ها خیلی به یکدیگر نزدیک باشد، از نقاط غالب مرتبه بعد استفاده می‌شود تا مراکز هر خوشه مشخص گردد که این نقاط در شکل (۶) نمایش داده‌شده است. سپس خوشه‌بندی با یکی از روش‌های خوشه‌بندی فازی C-میانگین یا خوشه‌بندی گاستافسون-کِسِل انجام می‌شود.



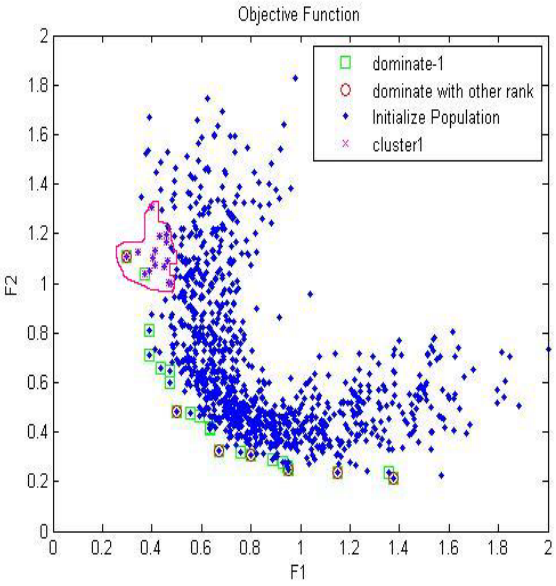
شکل (۵): مقادیر غالب امیدوارکننده



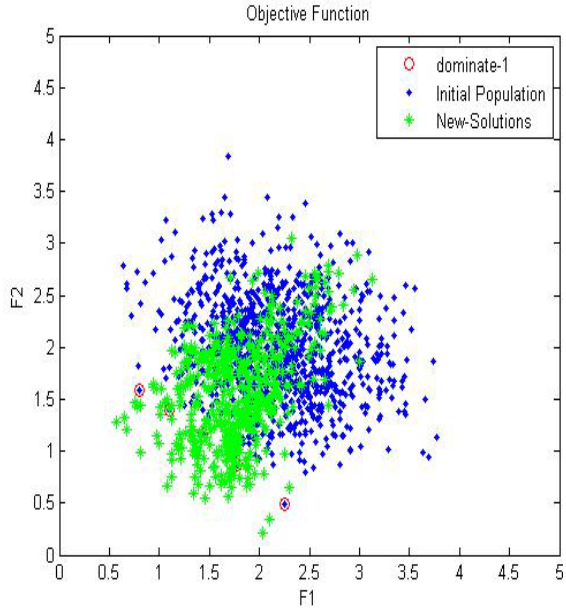
شکل (۶): تعیین مراکز خوشه‌ها

حال اعضای هر یک از خوشه‌ها به‌عنوان والدین در نظر گرفته می‌شوند و برای هر خوشه با استفاده از روش تحلیل مؤلفه‌های اساسی، مدل‌سازی انجام می‌شود. شکل (۷) اعضای یکی از خوشه‌ها که به روش خوشه‌بندی C-میانگین فازی انجام‌شده است را نشان می‌دهد.

سپس با استفاده از مقادیر ویژه و بردارهای ویژه که از روش تحلیل مؤلفه‌های اساسی به‌دست‌آمده است، داده‌های جدید ایجاد می‌شوند. بدین‌صورت که با به‌کارگیری توزیع نرمال گوسین، داده‌های جدید برای هر خوشه متناسب با بزرگی و جهت داده‌های هر خوشه، ایجاد می‌شوند. شکل (۸) داده‌های جدید (فرزندان) ایجادشده برای خوشه موجود در شکل (۷) (والدین)را نمایش می‌دهد.



شکل (۷): تعیین مراکز خوشه‌ها



شکل (۸): راه‌حل‌های جدید ایجادشده برای خوشه شکل (۷)

حال از جمعیت انتخابی اولیه و راه‌حل‌های جدید ایجادشده برای هر خوشه، جمعیت به‌اندازه N، بر اساس رتبه‌بندی و فاصله ازدحام ایجاد می‌شوند. حال شرط توقف بررسی‌شده، اگر شرایط نهایی حاصل شد متوقف می‌شوند، در غیر این صورت به مرحله انتخاب راه‌حل‌های کاندید امیدوارکننده بازمی‌گردند.

۴- ارزیابی کارآیی

رویکرد حل مسئله یکی از اساسی‌ترین بخش‌های یک پژوهش است که این مهم به این بخش اختصاص‌یافته است. با توجه به تعریف و مدل ریاضی مساله این پژوهش، روش حلی که برای آن در نظر گرفته‌شده است، الگوریتم چند هدفه تکاملی بر پایه مدل‌سازی معکوس با استفاده از فرآیند گوسین می‌باشد که به‌منظور روش حل مساله پیشنهادشده است.

طی سه دهه گذشته، تعداد زیادی الگوریتم تکاملی برای حل مسائل بهینه‌سازی چند هدفه ایجادشده است. با این وجود، عدم وجود یک بستر نرم‌افزاری به‌روز و جامع برای پژوهشگران برای ارزیابی صحیح الگوریتم‌های موجود و برای استفاده پژوهشگران از الگوریتم‌های انتخاب‌شده، برای حل مسائل می‌تواند به‌عنوان یک چالش مطرح باشد. وقتی کد منبع بسیاری از الگوریتم‌های پیشنهادی در دسترس عموم قرار نگرفته باشد، به‌منظور عدم مواجه با چنین چالشی، در این تحقیق، از یک سَکّو[1] در بستر نرم‌افزار متلب[2] برای بهینه‌سازی چند هدفه تکاملی، به نام پِلَت ای‌ام آ[3]، استفاده‌شده است که شامل بیش از ۵۰ الگوریتم تکاملی چند هدفه و بیش از ۱۰۰ مسئله آزمون چند هدفه، همراه با چندین عملکرد پرکاربرد است. این سکو کاملاً متن‌باز است، به‌گونه‌ای که کاربران می‌توانند بر اساس آن، الگوریتم‌های جدیدی بسازند.

کد منبع پِلَت ای‌ام آ، در آدرس[۲۵] قابل‌دسترسی است:

این نرم‌افزار مبنای حل دقیق برای مدل پیشنهادی این پژوهش است و هدف اعتبارسنجی مدل است. در بخش یافته‌های تحقیق، مدل پیشنهادی با روش دقیق و با استفاده از این نرم‌افزار اعتبارسنجی خواهد شد.

این تحقیق ازنظر هدف توسعه‌ای می‌باشد، چون به دنبال یافتن روشی علمی مناسب برای حل یک مساله است. ازنظر ماهیت داده‌ها آمیخته (کمی و کیفی) و ازنظر روش گردآوردی داده‌ها نیز توصیفی است. روش تحقیق در این پژوهش بر اساس ماهیت و نحوه گردآوری داده‌های آن، توصیفی (موردی و زمینه‌ای) است. با رویکرد آمیخته (کمی و کیفی) است. چون قرار است در یک مورد خاص عمیقاً پژوهش به عمل آید.

در این پژوهش، در دامنه وسیعی، آزمایش‌هایی به‌منظور نشان دادن عملکرد مدل مفهومی پیشنهادی انجام‌شده است و روش پیشنهادی با یک سری از الگوریتم‌ها، مقایسه و ارزیابی‌شده‌اند. این آزمایش‌ها بر روی نمونه تست‌های مشخص انجام‌شده است. بدین منظور در ادامه این بخش، ابتدا معیارهای ارزیابی و شاخص‌های موردمطالعه معرفی‌شده‌اند. در بخش دوم، تنظیماتِ پارامترها در آزمایش‌ها را نشان می‌دهد. در بخش سوم، الگوریتم‌های مقایسه شده را نشان می‌دهد. بخش چهارم به معرفی توابع ارزیابی می‌پردازد. در بخشِ یافته‌های تحقیق، نتایج آزمایش‌ها نشان داده می‌شود و در بخش بحث و نتیجه‌گیری، بررسی الگوریتم‌های مقایسه و ارزیابی‌شده را نشان می‌دهد.

۴-۱- معیارهای ارزیابی و شاخص‌های موردمطالعه

یک روش ساده برای ارزیابی کیفیت مجموعه‌ی راه‌حل‌ها، شاخص‌های کیفیت[4] است. به‌طورکلی، شاخص‌های کیفیت به شش گروه زیر دسته‌بندی می‌شوند [۱۶]: ۱- شاخص کیفیت برای همگرایی[5] ۲- شاخص کیفیت برای انتشار[6] ۳- شاخص کیفیت برای یکنواختی[7] ۴- شاخص قدرتمندی[8] ۵- شاخص کیفیت برای انتشار و یکنواختی ۶- شاخص کیفیت برای چهار بخش اول.

دو گروه همگرایی در شاخص‌های کیفیت وجود دارد:

۱- ارزیابی رابطه‌ی غلبگی پرتو[9] بین راه‌حل‌ها یا مجموعه‌ها.

۲- ارزیابی فاصله‌ی یک مجموعه راه‌حل از پرتو نما.

شاخص فاصله نسلی[10]، میانگین مربع فاصله اقلیدسی مجموعه‌ی راه‌حل‌ها، نسبت به نزدیک‌ترین نقطه در پرتو نما را اندازه‌گیری می‌کند [۱۷]. کیفیت انتشار[11]، مربوط به پوشش ناحیه یک راه حل است. تنوع خالص[12]، عدم شباهت هر راه حل به بقیه‌ی راه‌حل‌ها را در یک مجموعه راه حل نشان می‌دهد [۱۸]. شاخص کیفیت برای یکنواختی، توزیع یکنواختی یک مجموعه از راه‌حل‌ها را ارزیابی می‌کند و تغییرات فاصله بین راه‌حل‌ها را اندازه‌گیری می‌کند، همانند فاصله‌گذاری[13] [۱۹]. شاخص قدرتمندی، یک راه حل نامغلوب متفاوت را به مجموعه راه‌حل‌ها اضافه می‌کند تا باعث پیشرفت ارزیابی شود. شاخص کیفیت برای انتشار و یکنواختی، به هم نزدیک می‌باشند و می‌توانند با یکدیگر استفاده شوند تا تنوع مجموعه راه‌حل‌ها را نشان دهند و به دو گروه دسته‌بندی می‌شوند:

[1] Platform
[2] Matlab
[3] PlatEMO

[4] Quality Indicator
[5] Convergence
[6] Spread
[7] Uniformity
[8] Cardinality
[9] Pareto dominance
[10] Generational Distance (GD)
[11] Spread quality
[12] Pure Diversity (PD
[13] spacing

شاخص‌های بر مبنای فاصله

شاخص‌های برپایِ تقسیم نواحی[1]

که یک ناحیه مخصوص را به تعداد زیادی سلول‌های هم‌اندازه پارتیشن‌بندی می‌کنند و سپس تعداد سلول‌هایی که دارای راه‌حل‌ها هستند را محاسبه می‌کنند. بعضی از آن‌ها، سلول‌ها را مانند توری‌هایی در نظر می‌گیرند که فضا را به تعدادِ زیادی فرا جعبه[2] پارتیشن‌بندی می‌کنند، همانند اندازه‌گیری تنوع[3] [۲۱]. شاخص کیفیت برای تمام جنبه‌ها، همگرایی، انتشار، یکنواختی و اساسی را پوشش می‌دهد و به دو گروه دسته‌بندی می‌شوند:

شاخص کیفیت بر مبنای فاصله که فاصله پرتو نما نسبت به مجموعه راه‌حل‌های در نظر گرفته‌شده را اندازه‌گیری می‌نماید، همانند فاصله نسلی معکوس [۲۱ و ۷] [4].

شاخص بر مبنای حجم[5] که اندازه‌ی حجم را اندازه‌گیری می‌نماید و به مجموعه راه‌حل‌های در نظر گرفته‌شده، اختصاص می‌یابد، همانند فرا حجم [۲۲][6].

جدول (۱) ارتباط بین معیارهای ارزیابی و گروه‌بندی آن‌ها را نشان می‌دهد.

جدول (۱): ارتباط بین معیارهای ارزیابی و دسته‌بندی آن‌ها

دسته‌بندی	معیارهای ارزیابی					
	فرا حجم	فاصله نسلی معکوس	اندازه‌گیری تنوع	فاصله نسلی	تنوع خالص	فاصله‌گذاری
همگرایی	✓	✓		✓		
تنوع	✓	✓	✓		✓	
قدرتمندی	✓					
یکنواختی			✓			✓

در این تحقیق، شاخص فرا حجم، برای اندازه‌گیری راندمان نمونه‌های تست، به‌کار گرفته‌شده است و ۵۰۰ نقطه به‌صورت یکنواخت توزیع‌شده در هر نمونه، از پرتو نما، انتخاب می‌شود. برای ارزیابی عملکرد الگوریتم بهینه‌سازی چند هدفه تکاملی پیشنهادی،از تنظیمات پارامتر به شرح زیر استفاده‌شده است:

[1] Region division based indicators
[2] Hyper Box
[3] Diversity Metric (DM)
[4] Inversed Generational Distance (IGD)
[5] Volume-based
[6] Hyper Volume (HV)

۴-۲- تنظیمات پارامترها

برای تمام الگوریتم‌های مورد مقایسه، تنظیمات پارامتری یکسانی از مسائل، به‌کار گرفته می‌شود. ۲۰ اجرای مستقل برای هر الگوریتم مقایسه‌ای بر روی هر یک از نمونه‌های تست، انجام می‌شود. شرط پایان برای هر الگوریتم، حداکثر ۱۰۰۰۰۰ ارزیابی برازش[7]، برای تمام نمونه‌های تست، در نظر گرفته می‌شود.

تست مجموع رتبه‌بندی ویلکسون[8]، برای مقایسه نتایج به‌دست‌آمده، در یک سطح اهمیت ۰/۰۵ از الگوریتم‌های آزمایش‌شده به‌کار می‌رود. در جدول نتایج، نتایجِ آماری نشان داده می‌شود و برای هر الگوریتم، میانگین مقادیر و انحراف معیار، به ترتیب در دو سطح، نشان داده می‌شود. در تست مجموع رتبه‌بندی ویلکسون، برچسب + در جلوی نتایج، بدان معنا است که الگوریتم مقایسه شده بهتر از الگوریتم پیشنهادی است و برچسب - بدان معنا است که الگوریتم پیشنهادی از الگوریتم مقایسه شده بهتر است و برچسب ≈، بدان معنا است که تفاوت آماری زیادی بین نتایج مقایسه شده و الگوریتم‌های پیشنهادی وجود ندارد و همچنین بهترین نتایج در هر نمونه آزمایش، بارنگ مشکی پررنگ، مشخص‌شده است.

۴-۳- الگوریتم‌های مقایسه شده

اولین الگوریتم مقایسه‌ای، الگوریتم تکاملی چند هدفه دومین الگوریتم مقایسه‌ای، الگوریتم ژنتیک مرتب‌سازی نامغلوب-۲[9] است [۲۳] دومین الگوریتم مقایسه‌ای، الگوریتم تخمین چگالی چند هدفه بر پایه مدل منظم است [۷] و سومین الگوریتم مقایسه‌ای، الگوریتم تکاملی چند هدفه بر پایه مدل‌سازی معکوس با استفاده از فرآیند گوسین است [۹]. قوی‌ترین الگوریتم‌های تکاملی به‌منظور مقایسه انتخاب‌شده‌اند. با توجه به اینکه الگوریتم پیشنهادی، زیرمجموعه روش‌های الگوریتم‌های تکاملی چند هدفه بر مبنای مدل می‌باشد، دو الگوریتم مقایسه‌ای از این گروه انتخاب‌شده‌اند (الگوریتم تخمین چگالی چند هدفه بر پایه مدل منظم، الگوریتم تکاملی چند هدفه بر پایه مدل‌سازی معکوس با استفاده از فرآیند گوسین). همچنین الگوریتم پیشنهادی با یکی از بهترین الگوریتم‌های تکاملی چند هدفه سنتی (الگوریتم ژنتیک مرتب‌سازی نامغلوب-۲) مقایسه شده است.

۴-۴- توابع ارزیابی

روش پیشنهادی بر روی توابع تستِ ارائه‌شده توسط دب، تایل،

[7] Fitness Evalution (FE)
[8] Wilcoxon rank sum test
[9] Non-dominated Sorting Genetic Algorithm (NSGA-II)

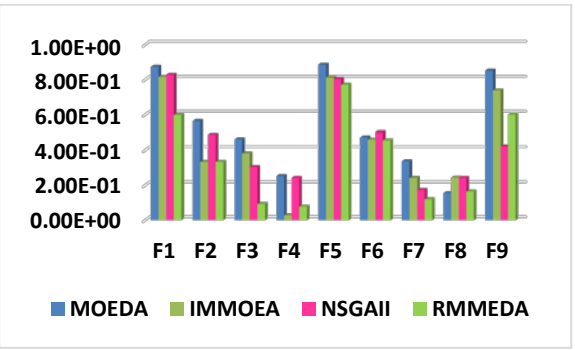
توماس (دی تی آل زِد)[1] آزمایش‌شده است[۷]. در این توابع تست، ارتباط بین متغیرهای تصمیم با استفاده از معادله (۶) می‌باشد:

$$\begin{cases} x_i \rightarrow \left(1+\alpha\frac{i}{n}\right)x_i \\ x_i^2 \rightarrow x_i^{\frac{1}{1+\beta\frac{i}{n}}} \end{cases} \quad (۶)$$

که در آن، i، اندیسِ هر متغیر تصمیم است و α و β، پارامترهای کنترلی هستند که $\alpha=3$ و $\beta=5$ است و تعداد متغیرهای تصمیم برابر با $n=30$ است.

۵- یافته‌های تحقیق

در این بخش، در این بخش، نتایج آزمایش‌ها نشان داده می‌شود. نتایج با الگوریتم‌های مقایسه شده، بر روی توابع تست که در بخش قبل معرفی‌شده، آزمایش می‌شود و با استفاده از معیارهای شاخص‌های ارزیابی بیان‌شده در بخش قبل، بر روی سَکّوِ پِلَت ای‌ام اُ ارزیابی‌شده‌اند و در جدول (۲) نشان داده‌شده است. از تست مجموع رتبه‌بندی ویلکسون، برای مقایسه نتایج به‌دست‌آمده، در یک سطح اهمیت ۰/۰۵ برای نمایش نتایج ارزیابی الگوریتم‌های آزمایش‌شده استفاده می‌شود. شکل (۹)، نتایج آماری بر روی الگوریتم‌های پیشنهادی و مقایسه شده، بر روی توابع تست را نشان می‌دهد.



شکل (۹): نتایج شاخصِ مقادیر فرا حجم بر روی تابع‌های تست به‌منظور ارزیابی الگوریتم‌های مقایسه‌ای و پیشنهادی

نتایج شکل (۹) بر اساس شاخص ارزیابی فرا حجم می‌باشد. علت استفاده از این شاخص آن است که هر سه معیار همگرایی، تنوع و قدرتمندی را به‌طور هم‌زمان مورد ارزیابی قرار می‌دهد.

بر اساس این شاخص که در جدول (۱) بیان‌شده است، معیارهای همگرایی، تنوع و قدرتمندی هر یک از الگوریتم‌های پیشنهادی و مقایسه‌ای مورد ارزیابی قرارگرفته و نتایج درج‌شده در شکل (۹)، درواقع برتری الگوریتم‌ها در هر یک از این معیارها به‌طور هم‌زمان را نشان می‌دهد. همچنین اعداد داخل پرانتز در هر خانه از جدول میزان انحراف معیار حاصل از ۲۰ اجرای مستقل را نشان می‌دهد.

در شکل (۹)، الگوریتم پیشنهادی بر روی توابع ارزیابی F_1,F_2,F_3,F_5,F_7,F_9 بهترین عملکرد را دارد و تابع F_4 بر روی الگوریتم تکاملی چند هدفه بر پایه مدل‌سازی معکوس با استفاده از فرآیند گوسین (IMMOEA) بهترین عملکرد را دارد و توابع F_6,F_8 بر روی الگوریتم ژنتیک مرتب‌سازی نامغلوب-۲ بهترین عملکرد را دارند.

با بررسی شکل (۹)، مشاهده می‌شود که الگوریتم پیشنهادی (MOEDA) در ۶ تابع ارزیابی، عملکرد بهتری نسبت به الگوریتم تکاملی چند هدفه بر پایه مدل‌سازی معکوس با استفاده از فرآیند گوسین (IMMOEA) دارد. همچنین الگوریتم تکاملی چند هدفه بر پایه مدل‌سازی معکوس با استفاده از فرآیند گوسین (IMMOEA) در ۴ مورد عملکرد بهتری نسبت به الگوریتم ژنتیک مرتب‌سازی نامغلوب-۲ (NSGA-II) دارد و الگوریتم ژنتیک مرتب‌سازی نامغلوب-۲ (NSGA-II) در ۷ مورد عملکرد بهتری نسبت به الگوریتم تخمین چگالی چند هدفه بر پایه مدل منظم (RMMEDA) دارد.

بنابراین طبق نتایج به‌دست‌آمده از شکل (۹)، مشاهده می‌شود که الگوریتم پیشنهادی عملکرد بهتری ازلحاظ همگرایی، تنوع و قدرتمندی نسبت به سایر الگوریتم‌های مقایسه‌ای دارد. درواقع وقتی عدد حاصل از ارزیابی فرا حجم در الگوریتم پیشنهادی بر روی توابع تست شده، نسبت به الگوریتم‌های مقایسه شده بیشتر است، به معنای کاهش تعداد ارزیابی و همگرایی بهتر در معیارهای همگرایی، تنوع و قدرتمندی می‌باشد. بنابراین روش پیشنهادی، الگوریتم تکاملی چند هدفه بر مبنای مدل را پیشنهاد می‌دهد که نمونه‌های ایجادشده از مدل، دارای تنوع، همگرایی و قدرتمندی بالاتری نسبت به روش‌های موجود می‌باشند و این امر به‌عنوان نقطه قوت الگوریتم پیشنهادی می‌باشد.

همچنین به‌کارگیری والدین به‌صورت خوشه‌های جداگانه و ایجاد مدل‌های جداگانه مبتنی بر هر خوشه، اثربخشی بهتری نسبت به سایر روش‌های بر پایه مدل دارد که در آن‌های مدل‌سازی بر روی تمامی فضای داده‌ها انجام می‌پذیرد. بنابراین می‌توان استنتاج نمود که نقطه قوت این روش به‌کارگیری مدل‌های مجزا مبتنی بر هر خوشه می‌باشد.

[1] Deb K, Thiele L, Laumanns M, Zitzler E (DTLZ)

۶- نتیجه‌گیری

در این پژوهش یک مدل مفهومی برای یک الگوریتم بهینه‌سازی چند هدفه تکاملی مبتنی بر مدل، با استفاده از خوشه‌بندی و تحلیل مؤلفه‌های اساسی، ارائه‌شده است. روش پیشنهادی با سایر الگوریتم‌های تکاملی چند هدفه سنتی که اخیر ارائه‌شده‌اند، مورد مقایسه و ارزیابی قرارگرفته شد و نتایج به‌دست‌آمده بر مبنای معیار ارزیابی فرا حجم می‌باشد که تمام دسته‌بندی‌های لازم (همگرایی، تنوع و قدرتمندی) را پوشش می‌دهد. با مقایسه نتایج ارائه‌شده در بخش یافته‌های تحقیق، مشاهده می‌شود که الگوریتم چند هدفه تکاملی بر پایه مدل‌سازی پیشنهادی، درمجموع عملکرد بهتری نسبت به سایر الگوریتم‌های مقایسه شده دارد و نرخ همگرایی، تنوع و قدرتمندی این الگوریتم نسبت به سایر الگوریتم‌های مقایسه شده بالاتر است. به‌کارگیری والدین به‌صورت خوشه‌های جداگانه و ایجاد مدل‌های جداگانه مبتنی بر هر خوشه، اثربخشی بهتری نسبت به سایر روش‌های بر پایه مدل دارد که در آن‌های مدل‌سازی بر روی تمامی فضای داده‌ها انجام می‌پذیرد. بنابراین می‌توان استنتاج نمود که نقطه قوت این روش به‌کارگیری مدل‌های مجزا مبتنی بر هر خوشه می‌باشد.

به‌عنوان پیشنهاد برای کارهای آتی می‌توان روش پیشنهادی را بر روی الگوریتم‌های تکاملی چند هدفه مبتنی بر مدل‌سازی معکوس با استفاده از فرآیند گوسین به‌کار برد. همچنین می‌توان روش پیشنهادی را در پیش‌بینی تأخیر در سیستم‌های کنترل مبتنی بر اینترن که از روش‌های فرا ابتکاری استفاده می‌کنند [۲۴]، به‌کار گرفت.

۷- مراجع

[1] H. Ma, H. Wei, Y. Tian, R. Cheng, and X. Zhang, "A multi-stage evolutionary algorithm for multi-objective optimization with complex constraints," Information Sciences, vol. 560, pp. 68-91, 2021.

[2] R. Cheng, C. He, Y. Jin, and X. Yao, "Model-based evolutionary algorithms: a short survey," Complex & Intelligent Systems, vol. 4, no. 4, pp. 283-292, 2018.

[3] M. Laumanns and J. Ocenasek, "Bayesian optimization algorithms for multi-objective optimization," in International Conference on Parallel Problem Solving from Nature, 2002, pp. 298-307: Springer.

[4] P. A. Bosman and D. Thierens, "Multi-objective Optimization with the Naive $$\mathbb {M} $$ ID $$\mathbb {E} $$ A," in Towards a New Evolutionary Computation: Springer, 2006, pp. 123-157.

[5] M. Pelikan, K. Sastry, and D. E. Goldberg, "Multiobjective hBOA, clustering, and scalability," in Proceedings of the 7th annual conference on Genetic and evolutionary computation, 2005, pp. 663-670.

[6] Y. Wang, J. Xiang, and Z. Cai, "A regularity model-based multiobjective estimation of distribution algorithm with reducing redundant cluster operator," Applied Soft Computing, vol. 12, no. 11, pp. 3526-3538, 2012.

[7] Q. Zhang, A. Zhou, and Y. Jin, "RM-MEDA: A regularity model-based multiobjective estimation of distribution algorithm," IEEE Transactions on Evolutionary Computation, vol. 12, no. 1, pp. 41-63, 2008.

[8] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," IEEE transactions on evolutionary computation, vol. 6, no. 2, pp. 182-197, 2002.

[9] R. Cheng, Y. Jin, K. Narukawa, and B. Sendhoff, "A multiobjective evolutionary algorithm using Gaussian process-based inverse modeling," IEEE Transactions on Evolutionary Computation, vol. 19, no. 6, pp. 838-856, 2015.

[10] B. Li, J. Li, K. Tang, and X. Yao, "Many-objective evolutionary algorithms: A survey," ACM Computing Surveys (CSUR), vol. 48, no. 1, pp. 1-35, 2015.

[11] P. D. Pantula, S. S. Miriyala, and K. Mitra, "An evolutionary neuro-fuzzy C-means clustering technique," Engineering Applications of Artificial Intelligence, vol. 89, p. 103435, 2020.

[12] I. Škrjanc, J. A. Iglesias, A. Sanchis, D. Leite, E. Lughofer, and F. Gomide, "Evolving fuzzy and neuro-fuzzy approaches in clustering, regression, identification, and classification: a survey," Information Sciences, vol. 490, pp. 344-368, 2019.

[13] R. Babuka, P. Van der Veen, and U. Kaymak, "Improved covariance estimation for Gustafson-Kessel clustering," in 2002 IEEE World Congress on Computational Intelligence. 2002 IEEE International Conference on Fuzzy Systems. FUZZ-IEEE'02. Proceedings (Cat. No. 02CH37291), 2002, vol. 2, pp. 1081-1085: IEEE.

[14] I. K. Fodor, "A survey of dimension reduction techniques," Lawrence Livermore National Lab., CA (US)2002.

[15] L. I. Smith, "A tutorial on principal components analysis," 2002.

[16] M. Li and X. Yao, "Quality Evaluation of Solution Sets in Multiobjective Optimisation: A Survey."

[17] D. A. Van Veldhuizen and G. B. Lamont, "Evolutionary computation and convergence to a pareto front," in Late breaking papers at the genetic programming 1998 conference, 1998, pp. 221-228.

[18] H. Wang, Y. Jin, and X. Yao, "Diversity assessment in many-objective optimization," IEEE transactions on cybernetics, vol. 47, no. 6, pp. 1510-1522, 2017.

[19] J. R. Schott, "Fault Tolerant Design Using Single and Multicriteria Genetic Algorithm Optimization," Air Force Inst Of Tech Wright-Patterson Afb Oh1995.

[20] A. Goli, H. K. Zare, R. Tavakkoli-Moghaddam, and A. Sadegheih, "Multiobjective fuzzy mathematical model for a financially constrained closed-loop supply chain with labor employment," Computational Intelligence, 2019.

[21] A. Zhou, Q. Zhang, Y. Jin, E. Tsang, and T. Okabe, "A model-based evolutionary algorithm for bi-objective optimization," in Evolutionary Computation, 2005. The 2005 IEEE Congress on, 2005, vol. 3, pp. 2568-2575: IEEE.

[22] K. Deb, Multi-objective optimization using evolutionary algorithms. John Wiley & Sons, 2001.

[23] W. Mkaouer et al., "Many-objective software remodularization using NSGA-III," ACM Transactions on Software Engineering and Methodology (TOSEM), vol. 24, no. 3, pp. 1-45, 2015.

[24] R. Asadzadeh Pourkarimi, R. Ghaffarpour , A.Khan Ahmadi," Delay Forecast in the Control System Based on Internet Using the Meta-Heuristic Methods and Comparing Methods with each Others ", Journal of “Innovations of Applied Information and Communication Technologies, vol. 1, no. 1, pp. 45-52, 1398.

[25]http://bimk.ahu.edu.cn/index.phps=/Index/Softwar/index.html.

Journal of "Innovations of Applied Information and Communication Technologies"
Vol. 1, No. 3, 2021 (Serial No. 3)

A Model-Based Evolutionary Algorithm using FCM-clustering and the PCA

Gh. Pezhman

Faculty of Computer Engineering and Information Technology, Shahid Sattari University of Aeronautical Sciences and Technology

Abstract

The structure of operators in most traditional multi-objective evolutionary algorithms are based on fixed heuristics such as intersection and mutation, so they are unable to learn the structures or properties of optimization problems. To equip the evolutionary algorithms with the learnability feature, model based evolutionary algorithms are presented. In the model evolutionary algorithms, innovative operators are replaced by machine learning models such as instructional and sample models. In this paper, a multi-objective model-based evolutionary algorithm is proposed in which in each generation, a probable area of the search space is modeled by a probabilistic model. In the decision area in the search space, which is comprised of the dominant points with better ranks, the clustering is performed either by different fuzzy methods, or by a crowding tournament selection operator on the first order dominant points. So, the points at close range to each other are removed and the result is considered to be the center of the clusters, and then, clustering is done based on the nearest neighbors. The principal component analysis algorithm, which is the best method for linearly reducing the given dimensions, is used for making the models. New solutions are obtained from the obtained model based on a normal distribution. The proposed method is tested and the results are compared with the method of non-dominated sorting genetic algorithms. The results show that this method is faster than earlier methods and has fewer repetitions and evaluation of functions, leading to better results.

Keywords: Crowding Tournament Selection Operator; Model-based evolutionary Algorithms; Fuzzy clustering; Multi-objective optimization; PCA.

* Corresponding author E-mail: pezhman.gholamnezhad@gmail.com