یک روش ترکیبی خوشه بندي مبتنی بر الگوریتم زنتیک با استفاده از عملگر های جدید تغییر

مصدر بیشمار، روابط سلطانیان و جوا نوری

چکیده:

مطالعه خوشه بندي به ميان مجموعه كردن مجموعه مجذر انحرافات را به دو نوع غیر خطي و غیر محدي
ورده و دراري تعداد روياب هاي ميلي است. هدف از این مقاله آنها روياب ترکيبي با استفاده از
K-Means الگوریتم زنتیک و براي خروج از نقاط ميلي است.

استفاده از الگوریتم زنتیک برای خروج از نقاط ميلي محلي توسيع محققتان بسياری انجام شده است.
در اين مقاله روي اين كنودري بر روی عملگرهای پاژكرتكي و ججش ارائه شده است. مبناد روست هاي
پاژكرتكي بر اين اثر است. استفاده از اين است. كه اگر عملگرهاي تغيير به جاي اگر بطور تصافدي در كلي نیاز
جاب عمال گرده، در يک منطقه محدود از پيش تعريف شده انجام شود، به جواب هاي پهري دست
خواهيم یافت.

براي ارزايي الگوریتم پاژكرتكي از هيه نوع عملگر ججش و پين نوع عملگر پاژكرتكي بر روی مجموعه
داده هاي استفاده هامان استفاده شده است. مقاييس نتایج بيدست آمده با سيبي رسش ها، به ازاي كلي
ميايونات مشاهده مي كنيد ميتوان با استفاده از عملگر پاژكرتكي ساده يك نقطه اي و عملگر ججش ارائه
شد. در اين مقاله با نام عملگر ججش منطقه اي خوشه بندي به جواب های پهري دست یافت.

1. مقدمه

خوشه بندي در فضال یک مجموعه داده يا نمونه به تعداد K گروه يا خوشه براساس شباهت
يا عدم شباهت آنها است. در واقع نمونه ها که در یک زمر
مجموعه قرار مي گيرند، بهم شبيه اند و يا آنهايي كه در زمر
مجموعه ديگر قرار مي گيرند متفاوت و غير شببه هستند.

تاريخ وصول: 89/09/29
تاريخ تصويب: 90/12/21

* گروه سيستم مسئول مقاله: دكتر مسعود بقيني، استاديار دانشگاه مهندسي
قوه آن، دانشگاه علم و صنعت ايران.
yaghini@iust.ac.ir
روایای سلطانیان، دانشجویی ارشد مدریت اداری، دانشگاه علم و
roya.soltanian@yahoo.com
صنعت ایران.
جوا نوری، دانشجویي كمپیوتر، دانشگاه علم و صنعت ایران.
javad.nouri@gmail.com

* Partitioning
1. Object
3. Mean value
که روش ترکیبی خوشه بندی مبتنی بر کرویت یکنیکی با استفاده از اقلیمی‌سی (که مزیت اصلی هر اقلیمیست) به کار رفته خوشه‌های مورد نیاز را مشخص می‌کند. این روش به دستور کارایی خوبی به درک اقلیمی‌سی در این روش را ارائه می‌دهند. 

2. تحقیقات انجام شده مرتبه

در دو دهه اخیر تحقیقات بسیاری در مورد روش‌های خوشه‌بندی تمایل رشدی داشته است. این تحقیقات به تشویق به اثبات گرفتن روش‌های خوشه‌بندی و در مورد تحقیقات اخیر در این زمینه، این روش‌ها به‌عنوان یک روش خوبی به‌کار رفته در این روش‌ها هستند.

1. Babu and Murty
2. Bit-string representation
3. K. Krishna and M. Narasimha Murthy
4. Gradient Descent Algorithm
5. S. Bandyopadhyay, U. Maulik
6. P. Hansen and N. Mladenovic
7. M. Laszlo and S. Mukherjee
8. Hyper-Quad Tree
9. Subdivision
10. {Sum of Squared Euclidean distance
11. Global Optimal

www.SID.ir
3. الگوریتم پیشنهادی

در روش پیشنهادی، از الگوریتم زنتیکی [16] به عنوان یک الگوریتم غیر منجرک شده که مقدار مراکز خوشه‌ها استفاده شده است. در این بخش تجزیه اجزای الگوریتم پیشنهادی می‌پردازیم.

3.1 جمعیت اولیه

الگوریتم پیشنهادی با مجموعه ای از کروموزوم‌ها که در کدام نشان دهنده یک جواب به سئوالی مفروض هستند، تحت عنوان جمعیت اولیه شروع به کار می‌کند. در تولید جمعیت اولیه در روش پیشنهادی این است که هر کروموزوم ویژه یک رشته بطول mxbه درون که در آن mرا برابر mاست مراکز خوشه‌ها است. بدین ترتیب از کروموزوم‌ها حتی در یک گاه‌های داده برگزی، جلوگیری نمی‌گردد و زمان اجرای الگوریتم وقت گیر و طولانی می‌شود.

3.2 عملکرد انتخاب

به منظور گزینش کروموزوم‌های والدین و وارد نمودن آنها به مرحله تولید مثل و تولید کروموزوم‌های جدید از عملکرد انتخاب کمک می‌گردد. در این روش دو کروموزوم انتخاب می‌شود که کمترین مقدار میزان پیش‌بینی کروموزوم منجرک کرده، سپس الگوریتم به مجموع SSE را به‌عنوان یک متغیر، مقدار ممکن مقداری روتوش می‌گیرد که حاصل این روتوش هر چه افزایشی کمتری باشد، شایستگی منجرک کرده. این انتخاب آن کروموزوم پیشتر می‌شود.

نخست در هر سل دو کروموزوم به مرحله تولید مثل می‌رسد و انتخاب هر بازیتی دیگر در هر سل با توجه به احتمال انتخاب هر کروموزوم، یک جفت کروموزوم انتخاب می‌شود و با رو به رو به مرحله تولید مثل الگوریتم دچار پیش‌بینی به همراه خیزش حرارتی می‌گردد. در این مقاله عملکرد انتخاب به کمک چرخ رولت[17] انجام شده است. برای برتری کردند مشکلات این روش ایجاد شایستگی ها می‌شود نبی، به دو روش مختلف سیگنال نشات آش‌بار به صحت مقیاس بندی اطمنان خارج حاصل شود.

1. Regional Crossover
2. D-X. Chang, X-D. Zhang, C-W. Zheng
3. Path-based Crossover
4. Krista Rizman Zalik
5. Jing Xiao, YuePing Yan, Jun Zhang and Yong Tang
6. The quantum-inspired genetic algorithm
7. Zhiven Yu and Hau-San Wong
8. Quantization -Based Clustering Algorithm (QBCA)
3-2 عاملگر پاره‌تکیبی

یکی از مهم‌ترین عاملگر‌هایی که در الگوریتم‌های قرار دادن هستند، عاملگر پاره‌تکیبی است که در طراحی و پیاده‌سازی جواب‌های حالی هدف به‌عنوان استعمال برای اعمال عاملگر پاره‌تکیبی، الگوریتم‌های مناسب مانند $P^1$ را می‌توان به پاره‌تکیبی در نظر گرفت. همچنین، $P^2$ یکی از الگوریتم‌های الگوریتم‌های پاره‌تکیبی است که به دلیل محدودیت و تعداد کمی از مقوله‌های استفاده در الگوریتم، به‌طور عمده در روش‌های متفاوت برای عاملگر پاره‌تکیبی استفاده می‌شود.

گرفته این‌ها می‌تواند منطقه نسبی از استفاده در الگوریتم‌های پاره‌تکیبی باشد. مطالعه به‌عنوان یکی از الگوریتم‌های پاره‌تکیبی در الگوریتم‌های پاره‌تکیبی استفاده شده است.

3-4 عاملگر جهش

در کل، الگوریتم عاملگر جهش در خلاصه استفاده در الگوریتم‌های پاره‌تکیبی است. اما در این روش این نتیجه در هنگامی که $R^1$ و $R^2$ انجام داده شود، به دست آمده است.

در نهایت، الگوریتم پاره‌تکیبی یکی از الگوریتم‌های پاره‌تکیبی است که به‌عنوان استفاده در الگوریتم‌های پاره‌تکیبی استفاده می‌شود.
جدول (1): معرفی مجموعه‌های مورد استفاده

<table>
<thead>
<tr>
<th>رفرنس</th>
<th>مرجع</th>
</tr>
</thead>
</table>

در ادامه نتایج آزمون‌های مدل بیشینه‌ای بر روی چند سری مجموعه‌ای که در چند مورد شرایط (1)، (2)، (3)، (4) مورد بررسی قرار گرفته بود، به خوبی مشخص گردید که این ابزار در نظر گرفتن هر سلسله‌ای از نتایج بهترین راه‌حل‌ها را به خوبی مشخص گردید. ضمناً، بهترین راه‌حل‌ها توسط ابزار برای این مجموعه‌ها که در اینجا نشان داده شده است، در قسمت‌های مختلف مالکیتی و مدیریت اجرایی کم‌ترین میزان طبقات کلیدی محسوب می‌شود. همچنین، جدول (2) نتایج آزمون‌ها را نشان می‌دهد.


gT D

4-1 نتایج آزمون‌ها برای مجموعه‌های مورد استفاده

<table>
<thead>
<tr>
<th>رفرنس</th>
<th>نتایج</th>
<th>محفظه</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1]</td>
<td>57</td>
<td>GTD</td>
</tr>
<tr>
<td>[2]</td>
<td>59</td>
<td>BPZ</td>
</tr>
<tr>
<td>[3]</td>
<td>62</td>
<td>IRIS</td>
</tr>
</tbody>
</table>

در نتیجه، مدل‌ها در کنار همکاری با همکاران، با قرار دادن بهترین راه‌حل‌ها به بیشینه‌ای بر روی مجموعه‌های مورد استفاده، می‌توانند بهترین راه‌حل‌ها را به خوبی مشخص گردی و در جواب‌های حاصل نتایج بیشینه‌ای بیشتری پیدا کنند.

5-1 جایگزینی

در این مرحله، بیشینه‌ای بیشینه‌ای بر روی مجموعه‌های مورد استفاده، می‌تواند بهترین راه‌حل‌ها را به خوبی مشخص گردی و در جواب‌های حاصل نتایج بیشینه‌ای بیشتری پیدا کند.

5-2 شرط خانه‌کردن

در این مرحله، بیشینه‌ای بیشینه‌ای بر روی مجموعه‌های مورد استفاده، می‌تواند بهترین راه‌حل‌ها را به خوبی مشخص گردی و در جواب‌های حاصل نتایج بیشینه‌ای بیشتری پیدا کند.

6-1 ارزیابی کیفیت خانه‌کردن

برای ارزیابی کیفیت خانه‌کردن، GTD، که در این مرحله، بهترین راه‌حل‌ها را به خوبی مشخص گردید. این نتایج نشان داده است که در جواب‌های حاصل نتایج بیشینه‌ای بیشتری پیدا کند.

6-2 شرط بیشینه‌ای

در این مرحله، بیشینه‌ای بیشینه‌ای بر روی مجموعه‌های مورد استفاده، می‌تواند بهترین راه‌حل‌ها را به خوبی مشخص گردی و در جواب‌های حاصل نتایج بیشینه‌ای بیشتری پیدا کند.

نتایج بین المللی مهندسی صنایع و مدیریت تولید، خرداد 1391-جلد 33-شماره 1

www.SID.ir
| جدول 2. مقایسه نتایج مدل پیشنهادی با الگوریتم های GTD مورد استفاده در تحقیقات پیشین برای مجموعه داده BPZ
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>تعادل</td>
<td>پیشنهادی</td>
<td>اختلاف</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

همانطور که مشاهده می‌شود، تفاوت تعادل از الگوریتم پیشنهادی به الگوریتم GTD به آیا $K$ متفاوت یکسان بوده و درصد اختلاف با آنها

| جدول 4. مقایسه نتایج مدل پیشنهادی با الگوریتم های IRIS مورد استفاده در تحقیقات پیشین برای مجموعه داده BPZ
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>روش حل</td>
<td>GA &amp; KGA Clustering</td>
<td></td>
</tr>
<tr>
<td>تعادل</td>
<td>پیشنهادی</td>
<td>اختلاف</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

همانطور که مشاهده می‌شود، تفاوت تعادل از الگوریتم پیشنهادی به الگوریتم BPZ به آیا $K$ متفاوت یکسان بوده و درصد اختلاف با آنها

| جدول 5. مقایسه نتایج مدل پیشنهادی با الگوریتم های GTD مورد استفاده در تحقیقات پیشین برای مجموعه داده BPZ
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>تعادل</td>
<td>پیشنهادی</td>
<td>اختلاف</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

همانطور که مشاهده می‌شود، تفاوت تعادل از الگوریتم پیشنهادی به الگوریتم BPZ به آیا $K$ متفاوت یکسان بوده و درصد اختلاف با آنها

نتیجه‌ی بنی‌المللی مهندسی صنایع و مدیریت تویله‌ام، خرداد 1391-جلد 33-شماره 1
5. نتیجه‌گیری
با توجه به اینکه در الگوریتم K-Means هنگام انتخاب می‌شود، ممکن است الگوریتم در دام بیشتری محل قرار گیرد و جواب بهتری را تولید ننماید. اگر جهت خروج از وضعیت بهبود محلی تریک الگوریتم فسوک ذاکر با الگوریتم زنبیک مدل خوشه‌بندی جدیدی ارائه شده است که سپر خروج از دام بهینه محلی تولید جواب بهتری می‌گردد.

در این مقاله سعی گردد به ارائه عملکرد بهتری برای الگوریتم زنبیک مشکل داده ایکان الگوریتم در نقطه بهینه محلی برطرف گردد الگوریتم پیشنهادی با استفاده از پنجره هایی بازتکیپی و سه عملکر جنسیت که مجموعاً 15 حالت مختلف را تشکیل می‌دهد، ارایه شده است.

نتایج بدست آمده نشان می‌دهد با استفاده از یک عملکر بازتکیپی ساده یک نقطه ای و عملکر جنسی متنوع یا خوشه ای می‌توان تئوری بهتری ساخته با سایر روش‌ها، بدست ایجاد واقع در این روش جنسی بازبندی داده ایکان مرکز خوشه مورد نظر بصورت قطعی از کل فضای جواب صورت نمی‌گیرد، بلکه مرکز خوشه مورد نظر با یک نقطه ای فضای تعیین یافته در خوشه مورد نظر جایگذاشته می‌شود.

مراجع
A Hybrid Clustering Method Using Genetic Algorithm with New Variation Operators

M.Yaghini*, R.Soltanian & J.Noori

Masood Yaghini, Assistance Professor of Railway Eng-Iran University of Science and Technology
Roya Soltanian, MSc student of Master of Business Administration-Iran University of Science and Technology
Javad Noori, BSc student of Computer Eng-Iran University of Science and Technology

Keywords
Clustering, Genetic algorithm, K-Means algorithm

ABSTRACT

The clustering problem under the criterion of minimum sum of squares is a non-convex and non-linear program, which possesses many locally optimal values, resulting that its solution often being stuck at locally optimal values and therefore cannot converge to global optima solution. In this paper, we introduce several new variation operators for the proposed hybrid genetic algorithm for the clustering problem. The novel mutation operator, called Clustering Regional Mutation, exchanges neighboring centers and a simple one-point crossover. The proposed algorithm identifies proper clustering. The experimental results are given to illustrate the effectiveness of the new genetic algorithm.

© 2012 IUST Publication, IJIEPM. Vol. 23, No. 1, All Rights Reserved