**فصل چهارم**

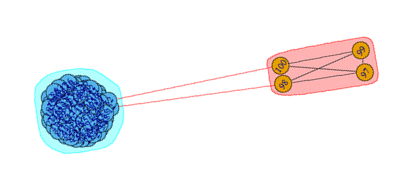
**نتایج پیاده سازی الگوریتم وارزیابی آن ها**

همان طور که قبلا ادعا شد ، استفاده از این رویکرد جدید برای انجمن یابی ، برخی از مشکلات شایع در مساله ی انجمن یابی را به خوبی مدیریت میکند. در این فصل به تفصیل ، این ادعا ها مطرح و نتایج بدست آمده بعد از پیاده سازی الگوریتم برای هر کدام از آنها بررسی و تحلیل می شود. در ادامه به ارزیابی نتایج این الگوریتم روی مجموعه داده های مصنوعی و مجموعه داده های معروف در حوزه انجمن یابی و مقایسه آن با نتایج چند الگوریتم معروف دیگر در این زمینه، پرداخته شده است.

**1.4 در ابتدا به بیان آزمایشها و بررسی نتایج روی شبکه های مصنوعی می پردازیم.**

**1.1.4آزمایش بررسی مساله محدودیت رزولوشن**

یکی از دلایل اصلی ارایه این رویکرد برای انجمن یابی ، برطرف کردن محدودیت مهم رزولوشن یا محدودیت اندازه انجمن های مختلف در شبکه ها بود. به این صورت که زمانی که اندازه یکی از انجمن ها بسیار بزرگ و دیگری بسیار کوچک باشد ، الگوریتم های انجمن یابی نمیتوانند ،انجمن کوچکتر را شناسایی کنند. همانطور که ادعا شد استفاده از اتاماتای سلولی یادگیرنده این محدودیت را برطرف می کند . برای نشان دادن این موضوع، عملکرد الگوریتم روی یک مجموعه داده ساختگی با شرایط زیر نشان داده شده است.

مجموعه داده ی زیر شامل 100 راس می باشد که از آن ها 96 راس در یک انجمن و با احتمال اتصال داخلی یک به همدیگر وصل هستند (گراف کامل 96 راسی) و4 تا راس باقی مانده در یک انجمن دیگر و با احتمال اتصال داخلی یک (گراف کامل4 راسی) و احتمال وجود یال بین این دو انجمن 0.01 می باشد. شکل 4- 1 نمایش گراف مورد نظر می باشد:

شکل 4- 1 : گراف شبیه سازی شده با 100راس که شامل زیرگراف کامل 96 و4 تایی است.

نتایج الگوریتم روی این گراف به صورت زیر است که در ادامه تحلیل آنها بیان شده است.

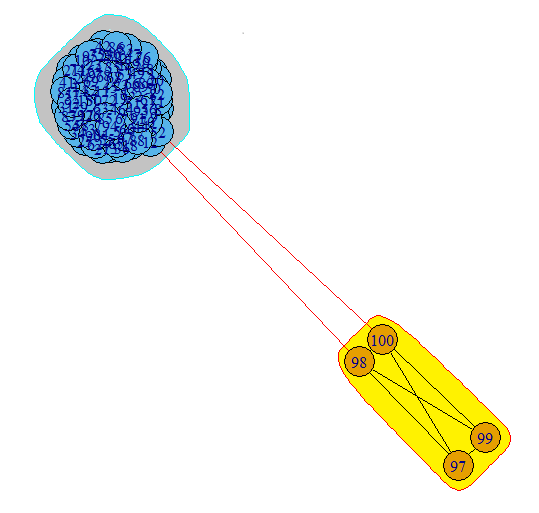
**Number of nodes: 100**

**Number of edges: 4568.0**

**Modularity is: 0.00262**

**NMI is: 1.0**

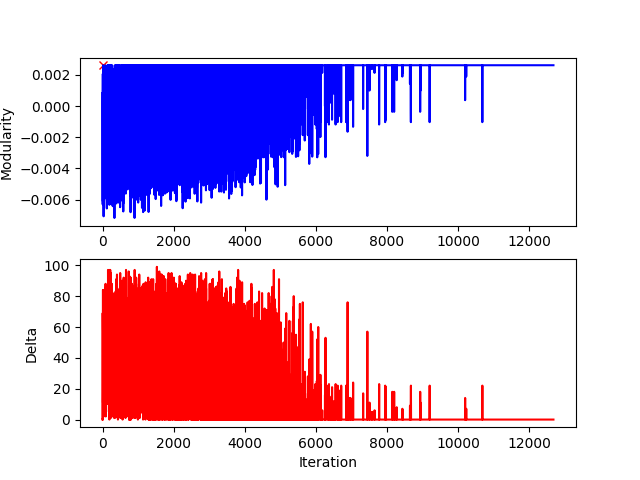
شکل 4-2 گراف بعد از اعمال الگوریتم CLA-NET نمایش داده شده است.



شکل 4-2 انجمن های پیدا شده روی گراف با الگوریتمCLA-net

همان طور که در شکل 4-3 گزارش شده است، الگوریتم به مقدار ماژولاریتی 0.00262همگرا شده و با تکرار انتخاب همسایگی جدید تغیری در نمودار عضویت پدیدار نمیشود لذا الگوریتم متوقف و مقدار نهایی ماژولاریتی را گزارش می دهد.

از آنجایی که گراف ساختگی می باشد،لذا برچسب های اصلی هر گره موجود است.حال برای ارزیابی عملکرد الگوریتم ،برچسب های اصلی با برچسب های پیدا شده توسط الگوریتم و توسط معیار NMI بررسی شده است. مقدار NMI=1 در این مثال نشان دهنده این است که بردار عضویت خروجی الگوریتم و بردار برچسب هایی گراف کاملا بر هم منطبق ولذا الگوریتم بدون مشکل توانسته انجمن های صحیح را تشخیص دهد و مشکل رزولوشن در این الگوریتم مطرح نیست.



شکل4-3 نمودار همگرایی ماژولاریتی و تغییرات دلتا در بررسی مساله محدودیت رزولوشن

**2.1.4آزمایش بررسی تاثیر نویز در عملکرد الگوریتم**

برای بررسی تاثیر نویز در مساله ، الگوریتم با مقدار اولیه alpharate=0.0003 و در حالت های مختلف زیر بررسی شده است.

ماتریس روبه رو مقدار ارتباط بین گراف ساختگی را نشان می دهد.به این صورت که درایه م این ماتریس نشان می دهد احتمال وجود یال بین انجمن و انجمن گروه چه مقداری است.

اندازه ماتریس همیشه مربعی و برابر m\*mمی باشد که mهمان تعداد انجمن های مورد نظر ما برای گراف ساختگی است.

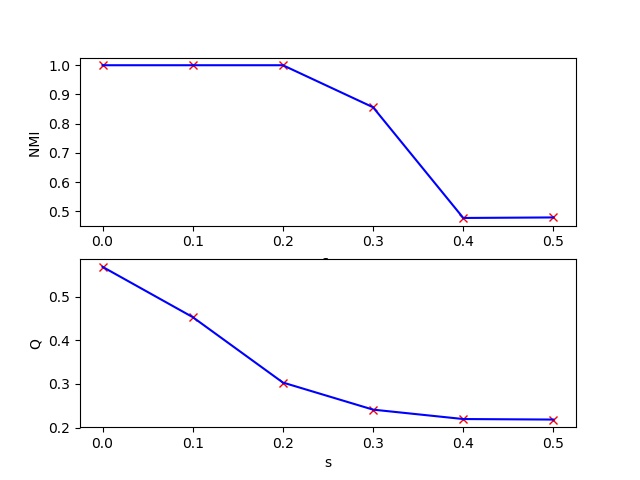
چنانچه مقدار sصفر باشد یعنی نویز وجود ندارد و انجمن ها کاملا مستقل از هم می باشند.

در اینجا برای مقدار های متفاوت sاز 0 تا 0.5 الگوریتم روی گراف ثابت 21 راسی با اندازه انجمن های به 8و 4 و 9 تایی ازمایش و نتایج به صورت زیر به دست امده است.برای هر مقدار به دلیل تصادفی بودن ده بار اجزا گرفته شده و میانگین مقدار ها در جدول مورد بررسی قرار گرفته است

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| S | دفعات اجرا | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Avg. |
| 0 | Q | 0.5681 | 0.5681 | 0.5681 | 0.5681 | 0.5681 | 0.5681 | 0.5681 | 0.5681 | 0.5681 | 0.5681 | **0.5681** |
| Iteration | 26910 | 32331 | 30129 | 27589 | 30903 | 26983 | 27812 | 29501 | 26501 | 29741 | **28840** |
| NMI | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | **1** |
| 0.1 | Q | 0.4530 | 0.4530 | 0.4530 | 0.4530 | 0.4530 | 0.4530 | 0.4530 | 0.4530 | 0.4530 | 0.4530 | **0.4530** |
| Iteration | 29622 | 29285 | 28850 | 28823 | 29688 | 30238 | 29255 | 30724 | 29214 | 29525 | **29522** |
| NMI | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | **1** |
| 0.2 | Q | 0.3033 | 0.3033 | 0.3033 | 0.3033 | 0.3033 | 0.3033 | 0.3033 | 0.3033 | 0.3033 | 0.3033 | **0.3033** |
| Iteration | 30535 | 32220 | 33106 | 32059 | 36292 | 31286 | 33296 | 31912 | 32277 | 34371 | **32725** |
| NMI | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | **1** |
| 0.3 | Q | 0.2413 | 0.2413 | 0.2413 | 0.2413 | 0.2413 | 0.2413 | 0.2413 | 0.2413 | 0.2413 | 0.2413 | **0.2413** |
| Iteration | 31895 | 33222 | 35317 | 33452 | 37505 | 34640 | 38316 | 36321 | 31592 | 32030 | **34429** |
| NMI | 0.856 | 0.856 | 0.856 | 0.856 | 0.856 | 0.856 | 0.856 | 0.856 | 0.856 | 0.856 | **0.856** |
| 0.4 | Q | 0.2198 | 0.2198 | 0.2198 | 0.2198 | 0.2198 | 0.2198 | 0.2198 | 0.2198 | 0.2198 | 0.2198 | **0.2198** |
| Iteration | 34240 | 35150 | 36384 | 34501 | 60921 | 33822 | 32655 | 39921 | 31794 | 82928 | **42231** |
| NMI | 0.4779 | 0.4779 | 0.4779 | 0.4779 | 0.4779 | 0.4779 | 0.4779 | 0.4779 | 0.4779 | 0.4779 | **0.4779** |
| 0.5 | Q | 0.2186 | 0.2186 | 0.2186 | 0.2186 | 0.2186 | 0.2077 | 0.2186 | 0.2186 | 0.2186 | 0.2186 | **0.2186** |
| Iteration | 37015 | 39838 | 34747 | 42451 | 36715 | 42329 | 36582 | 45204 | 40705 | 37649 | **39323** |
| NMI | 0.4788 | 0.4788 | 0.4788 | 0.4788 | 0.4788 | 0.4985 | 0.4788 | 0.4788 | 0.4788 | 0.4788 | **0.4788** |

|  |  |  |
| --- | --- | --- |
|  |  | 0 |
|  |  | 0.1 |
|  |  | 0.2 |
|  |  | 0.3 |
|  |  | 0.4 |
|  |  | 0.5 |

همانطورکه در شکل مشخص است با افزایش نویز احتمال خطا افزایش و لذا هردو معیار کاهش می یابند.اما حتی در بیشترین مقدار نویز ممکن هم معیار NMIاز 0.5 کمتر نشده واین پایداری الگوریتم را نسبت به نویز نشان می دهد.



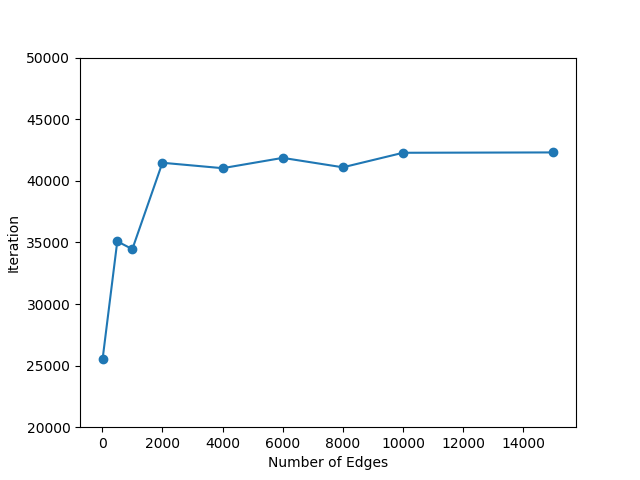
شکل 4-4 .نمودار تاثیرافزایش نویز S ، بر عملکرد الگوریتم با معیار هایNMI و ماژولاریتی

**3.1.4 آزمایش بررسی تاثیر اندازه شبکه در سرعت همگرایی الگوریتم**

آزمایش سوم طراحی شده تا سرعت همگرایی الگوریتم را بررسی کند. برای این منظور از مجموعه داده های ساختگی استفاده شد. از آنجایی که الگوریتم بر مبنای تعداد یالها و نه راس ها کار میکند،گراف های ساختگی باتعداد یال متفاوت اجرا و میانگین تعداد تکرارهای لازم برای همگرایی آن بررسی شد.

در این بررسی شبکه ها کاملا تصادفی و با میزان نویز های متفاوت انتخاب شده اند به همین منظور بهترین جواب لزوما1 NMI= نمی باشد. الگوریتم روی 9 گراف ساختگی تصادفی ،5 بار اجراشده است و نتایج آن در قالب جدول زیر بیان شده که ستون آخر بیانگر میانگین خروجی های الگوریتم است و نمودار تعداد تکرارها بر حسب یال ها در ادامه رسم شده است.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| تعداد  یالها |  | 1 | 2 | 3 | 4 | 5 | **Avg**. |
| 10  10 | Q | 0.4444 | 0.4444 | 0.4444 | 0.4444 | 0.4444 | **0.4444** |
| Iteration | 24963 | 29677 | 27643 | 22249 | 22952 | **25497** |
| NMI | 0.7489 | 0.7489 | 0.7489 | 0.7489 | 0.7489 | **0.7489** |
| 500  50 | Q | 0.3720 | 0.3720 | 0.3720 | 0.3720 | 0.3720 | **0.3720** |
| Iteration | 37609 | 35480 | 37676 | 36619 | 35080 | **35085** |
| NMI | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | **1.0** |
| 1000  75 | Q | 0.6259 | 0.6259 | 0.6259 | 0.6259 | 0.6259 | **0.6259** |
| Iteration | 32513 | 31334 | 39044 | 36764 | 32605 | **34452** |
| NMI | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | **1.0** |
| 2000  100 | Q | 0.3563 | 0.3563 | 0.3563 | 0.3563 | 0.3563 | **0.3563** |
| Iteration | 42868 | 40334 | 36052 | 42622 | 45445 | **41464** |
| NMI | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | **1.0** |
| 4000  150 | Q | 0.4298 | 0.4298 | 0.4298 | 0.4298 | 0.4298 | **0.4298** |
| Iteration | 41928 | 35080 | 39845 | 45215 | 43056 | **41024** |
| NMI | 0.7809 | 0.7809 | 0.7809 | 0.7809 | 0.7809 | **0.7809** |
| 6000  200 | Q | 0.2921 | 0.2921 | 0.2921 | 0.2921 | 0.2921 | **0.2921** |
| Iteration | 42473 | 41256 | 43356 | 39654 | 42569 | **41861** |
| NMI | .72590 | 7259.0 | 7259.0 | 7259.0 | 7259.0 | **7259.0** |
| 8000  225 | Q | 0.3603 | 0.3603 | 0.3603 | 0.3603 | 0.3603 | **0.3603** |
| Iteration | 41250 | 42045 | 45182 | 40952 | 36048 | **41095** |
| NMI | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | **1.0** |
| 10000  250 | Q | 0.4168 | 0.4168 | 0.4168 | 0.4168 | 0.4168 | **0.4168** |
| Iteration | 39938 | 42058 | 41256 | 43465 | 44659 | **42275** |
| NMI | .76600 | .76600 | .76600 | .76600 | .76600 | **.76600** |
| 15000  300 | Q | 0.3816 | 0.3816 | 0.3816 | 0.3816 | 0.3816 | **0.3816** |
| Iteration | 45647 | 42985 | 41595 | 39085 | 42202 | **42303** |
| NMI | 0.7604 | 0.7604 | 0.7604 | 0.7604 | 0.7604 | **0.7604** |

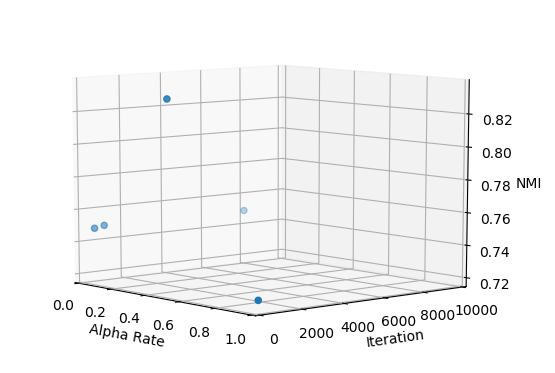


همانطور که در نمودار مشاهده میشود تعداد تکرارهای الگوریتم برای همگرا شدن بر حسب تعداد یال، نسبتا ثابت و در کل الگوریتم نسبت به افزایش یال ها ثبات خوبی نشان داده است.

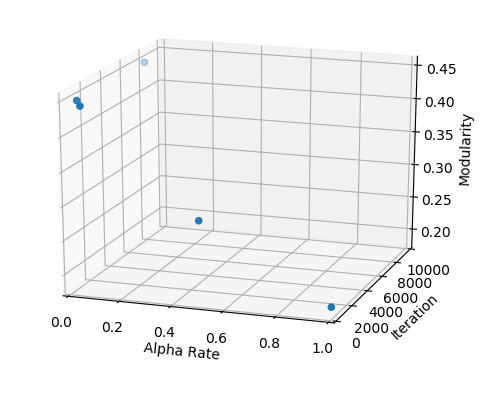
**4.1.4آزمایش بررسی تاثیر نرخ یادگیری در سرعت همگرایی**

نرخ یادگیری برای اتاماتاهای یادگیری در الگوریتم مقداری اختیاری و قابل تنظیم است. اما از انجایی که نرخ های بسیار پایین مانند alpharate=0.0003 باعث تعداد تکرار های بسیار بالای الگوریتم جهت همگرایی می شوند،لذا در این ازمایش روی یک گراف ثابت تصادفی الگوریتم با نرخ های متفاوت اجرا و نتایج مقایسه می شوند.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| نرخ یادگیری |  | 1 | 2 | 3 | **Avg**. |
| 1 | Q | 0.1234 | 0.4444 | 1.85 e-17 | **0.1892** |
| Iteration | 103 | 103 | 103 | **103** |
| NMI | 0.419 | 0.7489 | 1.0 | **0.7226** |
| 0.5 | Q | 0.4444 | 0.4444 | 1.85 e-17 | **0.2962** |
| Iteration | 108 | 108 | 103 | **106** |
| NMI | 0.7489 | 0.7489 | 1.0 | **0.8326** |
| 0.05 | Q | 0.4444 | 0.4444 | 0.4444 | **0.4444** |
| Iteration | 222 | 194 | 203 | **206** |
| NMI | 0.7489 | 0.7489 | 0.7489 | **0.7489** |
| 0.005 | Q | 0.4444 | 0.4444 | 0.4444 | **0.4444** |
| Iteration | 1170 | 937 | 918 | **1008** |
| NMI | 0.7489 | 0.7489 | 0.7489 | **0.7489** |
| 0.0005 | Q | 0.4444 | 0.4444 | 0.4444 | **0.4444** |
| Iteration | 7698 | 8423 | 7904 | **8008** |
| NMI | 0.7489 | 0.7489 | 0.7489 | **0.7489** |
| 0.00005 | Q | 0.4444 | 0.4444 | 0.4444 | **0.4444** |
| Iteration | 64706 | 67140 | 70483 | **67443** |
| NMI | 0.7489 | 0.7489 | 0.7489 | **0.7489** |



این نمودار نشان می دهد که در مقدار 0.5 برای نرخ یادگیری در این مثال خاص، بالاترین مقدار NMI با تقریبا کمترین تعداد تکرار الگوریتم به دست آمده است. از طرفی دیگر با رسم نمودار ماژولاریتی بر حسب نرخ یادگیری و تعداد تکرار مشاهده می شود که باز هم با افزایش مقدار نرخ یادگیری ماژولاریتی دیگر افزایشی نداشته است. لذا افزایش مقدار نرخ یادگیری به جز افزایش تعداد تکرار تاثیر دیگری ندارید.



**5.1.4آزمایش مقایسه نتایج الگوریتم با سایر الگوریتم های انجمن یابی**

**2.4 بررسی نتایج الگوریتم روی مجموعه داده های واقعی**