

تحلیل شبکه‌های زیستی علامت‌دار از دیدگاه نظریه تعادل ساختاری

مریم احسانی¹، محمد مهدی سپهری^{2*}، سید شهریار عرب³

1- دانشجوی دکتری مهندسی صنایع، دانشکده مهندسی صنایع و سیستم‌ها، دانشگاه تربیت مدرس، تهران

2- دانشیار، دانشکده مهندسی صنایع و سیستم‌ها، دانشگاه تربیت مدرس، تهران

3- استادیار، گروه بیوفیزیک، دانشکده علوم زیستی، دانشگاه تربیت مدرس، تهران

* تهران، کد پستی 1411713114

mehdi.sepohri@modares.ac.ir

(دریافت مقاله: 93/8/30 پذیرش مقاله: 93/11/25)

چکیده- تحلیل شبکه‌های زیستی علامت‌دار از زمینه‌های پژوهشی جدید در سال‌های اخیر به شمار می‌رود. این مقاله به بررسی و مقایسه خوشه‌های متعادل و انجمن‌های چگال به عنوان دو الگوی ساختاری متفاوت در تحلیل و کشف واحدهای ساختاری مؤثر در شبکه‌ها می‌پردازد. اگرچه بنابر مطالعات موجود شبکه‌های زیستی به سمت ساختار متعادل گرایش دارند، اما این پژوهش نشان می‌دهد که الگوی انجمن‌های چگال در شبکه بازنویسی سلول‌های *E.coli* از طرح خوشه‌های متعادل فاصله داشته و تعداد پیوندهای منفی در آنها به طور معناداری بالا است. این فاصله به عنوان معیاری در ارزیابی و تحلیل ساختار و عملکرد سیستم پیشنهاد می‌شود. همچنین با وجود آنکه بیشتر پیوندهای بین خوشه‌های متعادل از نوع منفی است، اما این تحقیق بر نقش پیوندهای مثبت واسطه بین خوشه‌های متعادل در پاسخ سیستم تأکید دارد. تحلیل داده‌های تعدادی زیر شبکه تنظیم ژنی از *E.coli* و *S.cerevisiae* نشان می‌دهد که اغتشاش در وضعیت ژن‌های این پیوندهای واسطه تأثیر مهمی در پاسخ کلی سیستم داشته و سبب ایجاد فاصله بیشتر تعادل جدید از وضعیت اولیه می‌شود. به این ترتیب شناسایی خوشه‌های متعادل و ارتباطات آنها، رویکرد مؤثری در تعیین واحدهای مؤثر عملکردی و ارتباطات کلیدی شبکه‌ها فراهم می‌کند. این رویکرد در کاربردهایی نظیر هدفگیری ژنی و قطع ارتباطات مؤثر بین زیر واحدهای اصلی سلولی می‌تواند مفید واقع شود.

کلیدواژگان: شبکه‌های علامت‌دار، خوشه متعادل، تعادل ساختاری، شبکه تنظیم ژنی.

1- مقدمه

مبانی اصلی دیدگاه زیست‌شناسی سیستمی است [1-3].
از همین‌رو شناسایی واحدها و الگوهای ساختاری نظیر

تأکید بر اهمیت تأثیر ساختار و ارتباطات و برهم‌کنش‌های
بین اجزا و پویایی آن بر عملکرد و رفتار کلی سیستم از

موتیف‌ها¹ و انجمن‌ها² از رویکردهای اصلی در تحلیل و بررسی شبکه‌های زیست مولکولی و تعیین زیرواحدهای اصلی سیستم به شمار می‌آید [4, 6].

در میان مدل‌های مختلف شبکه‌ای در تبیین سیستم‌های متنوع زیستی، گراف‌های علامت‌دار³ به منظور مدل‌سازی رده‌های گوناگونی از سیستم‌های زیستی با ملاحظه ارتباط مثبت یا منفی عناصر به کار می‌رود. به عنوان مثال می‌توان به شبکه‌های عصبی، شبکه‌های تنظیم ژنی⁴ و شبکه‌های پیام‌رسانی سلولی⁵ اشاره کرد که تأثیر عناصر آن‌ها بر یکدیگر از نوع فعال‌سازی (مثبت) یا بازدارنده (منفی) است. در این حوزه، نظریه سیستم‌های یکنوا⁶ [۸، ۷] و نظریه تعادل ساختاری⁷ [9] از تئوری‌های محوری در تحلیل ساختار و پویایی سیستم‌های با گراف علامت‌دار به شمار می‌روند.

یک شبکه علامت‌دار ساختار متعادل دارد، اگر هیچ حلقه‌ای با تعداد فرد یال منفی در آن وجود نداشته باشد. ثابت می‌شود که یک گراف دارای ساختار متعادل است اگر و فقط اگر همه ارتباطات آن مثبت باشند یا بتوان مجموعه رئوس آن را به دو بخش چنان افراز کرد که تمام ارتباطات بین عناصر هر بخش از نوع مثبت و تمام ارتباطات بین عناصر دو بخش از نوع منفی باشد [9].

Davis با کم کردن محدودیت‌ها تعادل ساختاری ضعیف را چنین تعریف می‌کند: یک گراف علامت‌دار دارای تعادل ساختاری ضعیف یا به عبارتی خوشه‌پذیر است اگر حلقه‌ای با دقیقاً یک یال منفی در آن وجود نداشته باشد. در این صورت ثابت می‌شود که مجموعه رئوس گراف را می‌توان به دو یا چند بخش چنان افراز کرد که همه ارتباطات بین عناصر هر بخش از نوع مثبت و هر ارتباط

بین دو عنصر از دو بخش متفاوت از نوع منفی باشد [10].

این فرضیه که شبکه‌های زیستی با گراف علامت‌دار در یک تکامل ساختاری به سمت تعادل ساختاری نزدیک می‌شوند یا شبکه‌های زیستی به تعادل ساختاری نزدیکند، در بعضی از پژوهش‌های اخیر مطرح شده است [8]. چنین ساختاری به عنوان یک ساختار نزدیک به یکنوا⁸، سبب پایداری بالا و رفتار تقریباً منظم و دور از آشوب در سیستم می‌شود. Iacono و همکاران، همچنین Lin و همکاران در پژوهش‌هایی میزان تعادل ساختاری در انواعی از شبکه‌های زیست مولکولی را بررسی کرده و نشان داده‌اند که در الگوی ساختاری این شبکه‌ها اغلب پیوندهای منفی ارتباط بین خوشه‌های با ارتباطات درونی غالباً مثبت را فراهم می‌کنند [۱۲، ۱۱]. این خوشه‌ها خوشه‌های متعادل⁹ نامیده می‌شوند.

در این مقاله طرح خوشه‌های متعادل و انجمن‌های چگال¹⁰ در تعدادی از شبکه‌های تنظیم بازنویسی سلول *Escherichia coli* تحلیل و با هم مقایسه شده است. میزان عدم تطابق بین خوشه‌های متعادل و انجمن‌های چگال به عنوان یک شاخص در ارزیابی عملکرد سیستم به ویژه ارزیابی رفتار سیستم در انتشار اطلاعات و اغتشاش معرفی می‌شود.

افراز شبکه بر مبنای تعادل ساختاری بخش‌هایی را شناسایی می‌کند که به عنوان زیر واحدهای نزدیک به یکنوا در تعیین عملکرد سیستم نقش دارند و از نتیجه افراز با رویکردهای دیگر مثلاً افراز بر مبنای چگالی متفاوتند. نقش پیوندهای منفی واسطه بین این بخش‌ها به عنوان پل¹¹ در پژوهش‌هایی مانند مقاله Lin و همکاران بررسی شده است. اما در این مقاله به تحلیل و بررسی

1. Motif
2. Community
3. Signed Graph
4. Grn: Gene Regulatory Network
5. Cell Signaling Network
6. Monotone Systems Theory
7. Structural Balance

8. Near-Monotone

9. Balanced Cluster

10. Density Based Community

11. Bridge

نقش پیوندهای مثبت در تعادل کلی سیستم پرداخته شده است. در واقع این پیوندها در پایداری و انتشار اطلاعات در سیستم نقش مهمی دارند و این واقعیت روی داده‌های تعدادی زیر شبکه از شبکه تنظیم بازنویسی¹² از سلول *E.coli* از سایت gnw.sourceforge.net نشان داده شده است. به این ترتیب از این روش می‌توان برای شناسایی ژن‌هایی که در تعادل سیستم نقش اصلی‌تری دارند، بهره برد. این رویکرد در کاربردهایی نظیر هدفگیری ژنی¹³ و قطع ارتباطات مؤثر بین زیرواحدهای اصلی سلولی می‌تواند مفید واقع شود.

در ادامه ابتدا در بخش بعد نظریه‌ها و مفاهیم پایه مورد نیاز مرور خواهند شد. سپس در بخش ارائه و تحلیل نتایج به تحلیل دو الگوی خوشه‌بندی متعادل و افراز بر مبنای چگالی در سلول *E.coli* پرداخته می‌شود. همچنین نقش ارتباطات مثبت بین خوشه‌های متعادل در پاسخ نهایی سیستم در زیرشبکه‌هایی از سلول *E.coli* و *Saccharomyces cerevisiae* تحلیل و بررسی شده است. در آخر پیشنهاداتی برای پژوهش‌های بعدی ارائه شده است.

2- مرور نظریه‌ها و مفاهیم پایه

2-1- نظریه تعادل ساختاری و سیستم‌های یکنوا

در بسیاری از مدل‌های شبکه‌ای، پیوندها به مفهوم مطلق ارتباط اشاره دارند. مثلاً ارتباط دوستی در یک شبکه اجتماعی یا جاذبه بین اجرام در یک سیستم مکانیکی. با وجود این مدل‌های دیگری وجود دارند که در آنها نوع مثبت یا منفی تأثیر در ارتباط دو عنصر اهمیت دارد، مانند دوستی یا دشمنی در یک شبکه اجتماعی، جاذبه یا دافعه در سیستم بارهای الکتریکی و اثر فعال سازی یا بازدارندگی در شبکه عصبی یا شبکه تنظیم ژنی. این

مفهوم با علامت‌گذاری پیوند در مدل گراف علامت‌دار نمایش داده می‌شود. برای یک خوشه سه‌تایی¹⁴ در یک گراف علامت‌دار، ارتباطات دو به دو سه راس به چهار صورت ممکن است (شکل 1).

ایده اصلی نظریه تعادل ساختاری ریشه در مطالعات‌هایدر حوزه روانشناسی دارد: هر ارتباط سه تایی که حاصل ضرب علامت‌ها منفی باشد، چالشی و نامتعادل است [13]. به این ترتیب از ترکیبات ممکن در شکل یک، دو ترکیب (الف) و (ب) یعنی دو دوست که دشمن مشترکی دارند و ارتباط دوستی دو به دو بین سه نفر، یک ارتباط متعادل محسوب می‌شود. دو ترکیب (ج) و (د) چالشی و نامتعادل بوده و انرژی اضافه‌ای به سیستم تحمیل می‌کند. بنابراین سیستم در جستجوی تعادل یا حداقل انرژی، گرایش به اصلاح علامت ارتباطات دارد.

یک گراف علامت‌دار، متعادل ساختاری است اگر هر سه تایی آن متعادل باشد. Cartwright و Harary مبانی نظریه تعادل ساختاری را توسعه داده و تعدادی قضایای اصلی را به زبان گراف بیان نمودند [9]. قضیه اصلی تعادل ساختاری بیان می‌کند که در یک گراف متعادل ساختاری یا هیچ پیوند منفی وجود ندارد و یا می‌توان مجموعه رئوس آن را به دو بخش چنان افراز نمود که تمام ارتباطات درون هر بخش از نوع مثبت و هر ارتباط بین دو بخش از نوع منفی باشد. هم چنین ثابت می‌شود که این شرط معادل آن است که هیچ حلقه‌ای با تعداد فرد یال منفی در گراف وجود نداشته باشد. این مطلب برای گراف‌های جهت‌دار هم با در نظر گرفتن نیم حلقه¹⁵‌ها (نیم حلقه: همان حلقه با صرف‌نظر از جهت کمان‌ها) صادق است. بعدها Davis، به عنوان تعمیمی از نظریه فوق، تعریف جدید تعادل ساختاری ضعیف را با آزادسازی محدودیت ترکیب نوع (د) ارائه داد.

14. Triad

15. Semi-Cycle

12. Transcription Regulatory Network

13. Gene Targeting



شکل 1 ترکیبات ممکن علامت ارتباطات در یک تریاد

اسپین³ $\sum$ به هر گره در گراف اسپین معینی را تخصیص می‌دهد. یک پیوند خشنود⁴ نامیده می‌شود اگر اسپین دو سر آن هم علامت و خود پیوند دارای علامت مثبت باشد یا اینکه علامت اسپین دو سر آن مخالف و پیوند دارای علامت منفی باشد. در غیر اینصورت پیوند ناخشنود⁵ یا ناسازگار در نظر گرفته می‌شود. به عبارتی اگر اسپین دو سر پیوند a_{ij} با σ_i و σ_j نشان داده شود، پیوند خشنود یا سازگار است اگر $\sigma_i * \sigma_j * a_{ij} = 1$ و در غیر این صورت پیوند ناخشنود یا ناسازگار نامیده می‌شود (شکل 3).

یک گراف علامت‌دار نسبت به الگوی تخصیص اسپین $\sum$ سازگار⁶ نامیده می‌شود اگر به ازای آن هیچ پیوند ناسازگاری در گراف نتیجه نشود.

یک سیستم پویا با مدل گراف علامت‌دار یکنوا نامیده می‌شود اگر برای آن تخصیص اسپینی موجود باشد به نحوی که گراف آن سازگار باشد. سیستم‌های یکنوا رفتار منظم داشته و از پایداری و استحکام بالایی برخوردارند. به همین سبب این فرضیه مطرح می‌شود که سیستم‌های زیستی ساختاری نزدیک به یکنوا دارند [8]. می‌توان ثابت کرد که طرح تخصیص اسپین سازگار برای یک گراف وجود دارد اگر و فقط اگر گراف دارای ساختار متعادل باشد. به این ترتیب ارتباط نظریه تعادل ساختاری و سیستم‌های یکنوا روشن می‌شود.

در این حالت مجموعه عناصر گراف قابل افراز به دو یا چند بخش است به نحوی که ارتباطات درون هر بخش از نوع مثبت و هر ارتباط بین دو عنصر از دو بخش متفاوت از نوع منفی باشد. چنین گرافی را خوشه‌پذیر¹ می‌نامند [10]. ثابت می‌شود این شرط معادل آن است که هیچ حلقه‌ای با دقیقاً یک یال منفی در گراف وجود نداشته باشد (شکل 2).

این مفاهیم ملاحظات جدیدی را در مطالعه ساختار و تکامل شبکه‌های علامت‌دار مطرح می‌کند. به عنوان مثال یک شبکه از تعادل ساختاری به چه میزان فاصله دارد و این فاصله چه تأثیری در رفتار و خصوصیات سیستم دارد؟ آیا تکامل یک سیستم تحت تأثیر گرایش به تعادل ساختاری بیشتر قرار می‌گیرد؟ نقش و تأثیر خوشه‌های متعادل و ارتباطات بین آنها در عملکرد سیستم چگونه است؟ به این ترتیب خصوصیات ساختاری و قوانین پویایی و تغییرات ساختاری سیستم‌هایی با مدل گراف علامت‌دار در حوزه‌های گوناگونی از شبکه‌های طبیعی تا شبکه‌های اجتماعی مورد پژوهش می‌باشد، به عنوان مثال می‌توان از [۱۴، ۱۱-۱۸] نام برد.

نظریه دیگری که با مفاهیم تعادل ساختاری ارتباط نزدیکی دارد، نظریه سیستم‌های یکنوا در حوزه کنترل و سیستم‌های پویا است. فرض می‌شود هر گره در شبکه با یکی از دو وضعیت - که با 1 یا -1 نمایش داده می‌شود - متناظر است. برگرفته از واژگان فیزیک آماری این وضعیت را اسپین² می‌نامند. یک الگوی تخصیص اسپین³

3. Spin Assignment

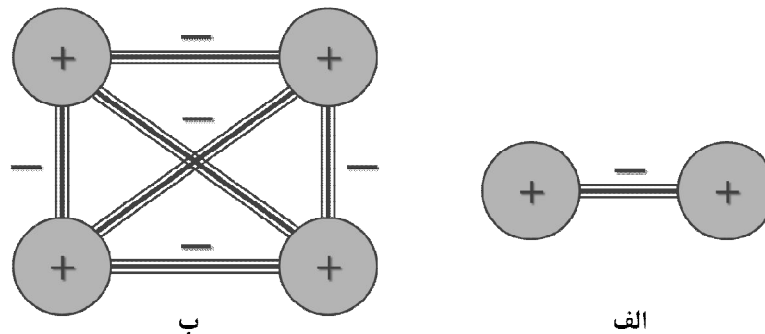
4. Satisfied

5. Frustrated

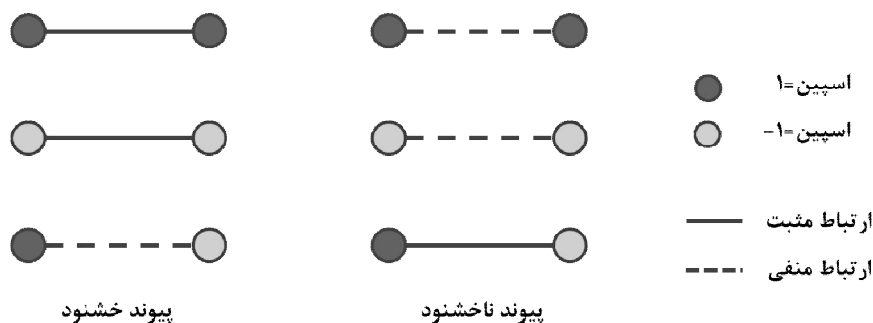
6. Consistent

1. Clusterable

2. Spin



شکل 2 الف - تعادل ساختاری در قضیه Harary و ب - تعمیم آن در مفهوم تعادل ساختاری ضعیف در تعریف Davis



شکل 3 الف - پیوندهای سازگار و ب - پیوندهای ناسازگار

ساختاری از حالت تعادل و دیگری فاصله طرح تخصیص اسپین از حالت بهینه، در واقع پیوندهای ناخشنود نوعی عدم تعادل را در سیستم به وجود می‌آورند که سیستم با حداقل نمودن تعداد آنها به دنبال حداقل کردن سطح انرژی خود است. به این ترتیب، پویایی سیستم به مفهوم تغییرات اسپین گره‌ها به واسطه تعامل با یکدیگر در طول زمان و تکامل ساختاری شبکه به مفهوم اصلاح پیوندها و علامت آن‌ها، هر دو را می‌توان در این راستا تعبیر کرد.

2-2- انجمن‌های چگال و خوشه‌های متعادل

بسیاری از شبکه‌های پیچیده واقعی ساختار خوشه‌پذیر دارند به این مفهوم که متشکل از گروه‌هایی هستند که ارتباطات درونی زیاد و ارتباطات بینابینی کم دارند. چنین ساختاری در انواع شبکه‌های اجتماعی، زیستی، کامپیوتری و غیره دیده می‌شود. به بیان ساده، انجمن‌ها در یک گراف به گروه‌های چگال که ارتباطات درونی زیاد و ارتباط

در این شرایط گراف بنابر قضیه تعادل ساختاری به دو بخش قابل افراز است. تنها دو طرح اسپین سازگار برای چنین گرافی وجود دارد: اسپین مثبت برای یک بخش و اسپین منفی برای دیگری و برعکس.

متداول‌ترین معیاری که برای ارزیابی فاصله یک گراف از تعادل ساختاری ارائه شده است، شاخص ناخشنودی¹ است. این شاخص برابر حداقل تعداد پیوندهای ناسازگار ممکن در یک گراف به ازای تخصیص اسپین‌های متفاوت است. برای یک گراف متعادل به ازای طرح‌های اسپین یاد شده در بالا همه پیوندها سازگارند و بنابراین مقدار این شاخص برابر صفر است. اما حتی برای چنین گرافی به ازای سایر طرح‌های تخصیص اسپین، تعدادی از پیوندها ناسازگار خواهند بود. به این ترتیب مشاهده می‌شود که ناسازگاری در گراف حاصل دو مؤلفه است: یکی فاصله

1. Frustration Index

بینابینی کم دارند، اشاره دارند. شناسایی و تحلیل این گروه‌ها و عملکرد و ارتباط آن‌ها به عنوان واحدهای اصلی ساختاری در تحلیل سیستم کلی از اهمیت خاص برخوردار است.

در افزاز گراف به انجمن‌ها اصولاً علامت پیوندها در نظر گرفته نمی‌شود. الگوریتم‌های متعددی به منظور کشف انجمن‌ها در گراف بر مبنای معیارها و محدودیت‌های متفاوت ارائه شده است. الگوریتم Girvan-Newman با حذف پیوندهای با بزرگترین مقدار میانداری¹ در یک رویه بازگشتی به شناسایی انجمن‌ها در گراف می‌پردازد [6]. در رویکرد دیگری الگوریتم Louvain تابع هدف بیشینه‌سازی مقدار مدولاریتی² را برای شناسایی بهترین ترکیب انجمن‌ها معرفی می‌کند [19]. مدولاریتی اشاره به میزان تفکیک پذیری ساختار به گروه‌ها دارد و برای یک گروه‌بندی، برابر تفاضل نسبت پیوندهایی که درون گروه‌ها قرار می‌گیرد و مقدار امید این نسبت چنانچه پیوندها به صورت تصادفی در گراف توزیع شوند، تعریف می‌شود. بنابراین مدولاریتی معیاری برای اندازه‌گیری کیفیت گروه‌بندی به عنوان گروه‌های چگال با ارتباطات تنک³ بین گروهی می‌باشد.

در افزاز گراف‌های علامت‌دار رویکرد دیگری بر مبنای تعریف خوشه در تعادل ساختاری ضعیف مطرح می‌شود. هدف افزاز در این رویکرد، بخش‌بندی گراف به خوشه‌هایی است که ارتباطات درون آن‌ها تا حد امکان از نوع مثبت و ارتباطات بین خوشه‌ها تا حد امکان از نوع منفی باشد. در یک گراف متعادل ساختاری یا خوشه پذیر در تعریف Davis، تعیین این خوشه‌ها به نحوی که همه ارتباطات درونی مثبت و همه ارتباطات بینابینی منفی باشد، امکان پذیر است. در این حالت در این خوشه‌بندی، تعداد یال‌های خطا یعنی یال‌های منفی درون خوشه‌ها و

یال‌های مثبت بین خوشه‌ها صفر است. تقریباً هیچ شبکه واقعی که این شرایط به طور کامل در آن برقرار باشد، موجود نیست. در عین حال شبکه‌های واقعی به چنین ساختاری نزدیکند. هرچه فاصله گراف از حالت خوشه‌پذیری بیشتر باشد، تعداد یال‌های خطا در این خوشه‌بندی بیشتر می‌شود. Doreian و همکاران الگوریتم‌هایی برای جستجوی خوشه‌بندی با مفهوم فوق ارائه داده‌اند [20، 21].

3- نتایج و تحلیل

در این پژوهش از داده شبکه تنظیم بازنویسی سلول *E.coli* از سایت RegulonDb در آدرس <http://regulondb.ccg.unam.mx/> استفاده شده است. این داده گراف علامت‌دار و جهت‌دار شبکه را فراهم می‌کند. بزرگترین مؤلفه همبند ضعیف این شبکه 1308 گره دارد که آن را *net a* می‌نامیم. تعدادی از گره‌ها در این گراف جهت‌دار روی هیچ عنصر دیگری تأثیر ندارند. بنابراین با حذف این گره‌ها که درجه خروجی آن‌ها برابر صفر است گرافی با 155 گره بدست می‌آید که بزرگترین مؤلفه همبند ضعیف آن 85 گره دارد که آن را *net b* می‌نامند. در این بخش به ارائه و تحلیل نتایج الگوریتم‌های شناسایی انجمن‌های چگال و خوشه‌های متعادل در این گراف و همچنین زیرگراف‌های مربوط به شبکه تنظیم ژنی از سلول *E.coli* و *S.cerevisiae* از سایت gnw.sourceforge.net پرداخته می‌شود.

3-1- انجمن‌های چگال و خوشه‌های متعادل: تطبیق دو الگو

برای گراف‌های *net a* و *net b* تعیین انجمن‌های چگال با استفاده از نرم‌افزار Pajek 3.13 با الگوریتم Louvain و خوشه‌بندی متعادل بر مبنای الگوریتم‌های نرم‌افزار Pajek 3.13 انجام شده است. چنانکه شرح داده شد، انجمن‌ها

1. Betweenness
2. Modularity
3. Sparse

ورودی³ می‌دهند که در آن یک فاکتور بازنویسی گروهی از ژن‌ها را به صورت بازدارنده کنترل می‌کند.

جدول 1 نتایج خلاصه آشکارسازی انجمن‌ها در net a. علامت ستاره ساختار موتیف SIM را نشان می‌دهد. چگالی کلی شبکه برابر 0/0015 و نسبت پیوندهای منفی برابر 0/47 می‌باشد.

شماره انجمن	رتبه (تعداد راس)	تعداد کمان	نسبت تعداد کمان‌های منفی	چگالی انجمن
1	47	93	0/1398	0/086
2	102	124	0/5323	0/0241
3	89	136	0/1958	0/0347
*4	11	10	1	0/1818
5	239	520	0/4769	0/0183
6	271	441	0/3719	0/0121
7	113	135	0/5333	0/0213
8	86	108	0/3889	0/0295
9	24	32	0/1875	0/01159
10	121	152	0/6579	0/0209
11	69	94	0/5745	0/0401
12	80	97	0/8969	0/0401
13	19	18	0/2222	0/1053
*14	4	3	1	0/5
*15	5	4	1	0/4
*16	19	21	0/9048	0/1228
*17	9	8	1	0/2222

تعیین انجمن‌ها در گراف net a. همچنین با مقدار پارامتر تفکیک‌پذیری برابر 0/3 در نرم افزار پاژک⁴ انجام شده است و تعداد انجمن‌ها برابر 7 و مقدار مدولاریتی برابر 0/76 به دست آمده است. اطلاعات مربوطه در جدول 2 آمده است. باز هم نسبت بالای تعداد یال‌های منفی در انجمن‌ها، گویای اختلاف بین دو الگوی افراز مزبور است.

به گروه‌های چگال با ارتباطات کم بینابین صرف‌نظر از علامت ارتباط اشاره دارند در حالی که خوشه‌های متعادل زیرگراف‌هایی هستند که ارتباطات درونی آن‌ها تا حد امکان از نوع مثبت و ارتباطات بین آن‌ها تا حد امکان از نوع منفی باشد. در اجرای الگوریتم مورد نظر پارامتر تفکیک‌پذیری¹، تعداد و دقت انجمن‌ها را کنترل می‌کند. مسأله محدودیت تفکیک‌پذیری² در موضوع شناسایی انجمن‌ها به این نکته اشاره می‌کند که در الگوریتم‌های شناسایی انجمن بر مبنای بهینه‌سازی مدولاریتی امکان تعیین انجمن‌هایی که اندازه آن‌ها از یک حد معین پایین‌تر است وجود ندارد [19]. بنابراین در اجرای الگوریتم مورد نظر اندازه کوچکترین انجمن‌ها با پارامتر تفکیک‌پذیری کنترل می‌شود. هرچه مقدار این پارامتر بزرگتر باشد تعداد بیشتری انجمن (در ابعاد کوچکتر) تمیز داده می‌شود.

تعداد انجمن‌ها در net a. (با مقدار پارامتر $res=0/8$) برابر 17 و مقدار مدولاریتی برابر 0/67 به دست آمده است. اطلاعات خلاصه آن در جدول 1 آورده شده است. اگر طرح انجمن‌های چگال به الگوی خوشه‌های متعادل نزدیک باشد، نسبت تعداد پیوندهای منفی در هر انجمن باید از میانگین این نسبت در کل شبکه بسیار کمتر باشد. اما نتایج نشان می‌دهد که در یازده تا از این انجمن‌ها این نسبت حتی از این مقدار بزرگتر است. از طرف دیگر نزدیکی دو الگو ایجاب می‌کند که چگالی خوشه‌های متعادل به چگالی متوسط انجمن‌ها نزدیک باشد، اما باز هم نتایج نشان می‌دهد که چگالی خوشه‌های متعادل از متوسط چگالی انجمن‌ها فاصله زیادی دارد. علاوه بر این تحلیل گراف نشان می‌دهد که هیچ حلقه جهت‌دار منفی در گراف وجود ندارد، با وجود این پیوندهای منفی در ساختار انجمن‌ها نقش مهمی دارند. در واقع بررسی ساختار آن‌ها نشان می‌دهد که پنج انجمن تماماً از پیوندهای منفی تشکیل شده‌اند و تشکیل یک موتیف تک

3. Sim: Single Input Module

4. Pajek

1. Resolution

2. Resolution Limit Problem

نتایج تحلیل مربوط روی net b. در جدول‌های 3 و 4 خلاصه شده است و همین نتیجه‌گیری را تأیید می‌کند.

جدول 2 نتایج خلاصه آشکارسازی انجمن‌ها در net a.

(resolution= 0.8, modularity= 0.67), علامت ستاره ساختار

موتیف SIM را نشان می‌دهد.

چگالی انجمن	نسبت تعداد کمان‌های منفی	تعداد کمان	مرتبه (تعداد راس)	شماره انجمن
0/0086	0/4134	537	353	2
0/0078	0/5332	1069	524	3
0/1159	0/1875	32	24	4
0/0342	0/8889	90	73	5
0/1228	0/9048	21	19	*6
0/2222	1	8	9	*7

جدول 3 نتایج خلاصه آشکارسازی انجمن‌ها در net b.

(resolution= 0.3, modularity= 0.74), علامت ستاره ساختار

موتیف SIM را نشان می‌دهد. چگالی کلی شبکه برابر 0/16 و نسبت پیوندهای منفی برابر 0/3934 می‌باشد.

چگالی انجمن	نسبت تعداد کمان‌های منفی	تعداد کمان	مرتبه (تعداد راس)	شماره انجمن
0/1345	0/3478	23	19	1
0/049	0/3827	81	58	2
0/25	0/4286	7	8	3

جدول 4 نتایج خلاصه خوشه‌بندی بر مبنای تعادل ساختاری

در شبکه net b.

چگالی انجمن	نسبت تعداد کمان‌های منفی	تعداد کمان	مرتبه (تعداد راس)	شماره انجمن
0/0526	0/1	10	20	1
0/0546	0/0426	47	42	2
0/0474	0	12	23	3

به طور خلاصه، نتایج نشان می‌دهد که اگرچه ساختار کلی به یک ساختار متعادل نزدیک است اما انجمن‌های مزبور تعداد بسیار بیشتری لینک‌های منفی نسبت به

وضعیت تعادل دارند. به بیان دیگر دو الگوی خوشه‌های متعادل و انجمن‌های چگال از یکدیگر فاصله دارند. نتایج همچنین نشان می‌دهد که این لینک‌های منفی هیچ حلقه منفی جهت‌داری را تشکیل نمی‌دهند و در عین حال در ایجاد موتیف‌های شبکه نقش اساسی دارند.

به این ترتیب مشاهده می‌شود که با وجود اینکه ساختار گراف‌های مزبور به تعادل ساختاری نزدیک است، اما دو الگوی افراز بر مبنای چگالی و افراز بر مبنای تعادل ساختاری از یکدیگر فاصله دارند و بنابراین واحدهای ساختاری متفاوتی را تعیین می‌کنند. چنانکه Sontag پیشنهاد می‌کند از آنجا که سیستم‌های زیستی از خود رفتار تقریباً منظم و با استحکام بالا نشان می‌دهند، منطقی است که از ساختار یکنوا یا حداقل نزدیک به یکنوا برخوردار باشند [8]. تعیین زیرواحدهای نزدیک به یکنوا در چنین سیستمی، واحدهای اساسی عملکردی را به دست می‌دهد. در پژوهش حاضر نشان داده شد که این واحدها با واحدهایی که با رویکرد متداول تعیین گروه‌ها بر مبنای چگالی به دست می‌آید، تفاوت دارند و بنابراین این روش می‌تواند به عنوان رویکرد متفاوتی در شناسایی واحدهای بیولوژیک مفید باشد.

نظریه سیستم‌های یکنوا و نزدیک به یکنوا ویژگی‌ها و رفتار سیستم‌هایی با گراف متعادل ساختاری یا نزدیک به آن را بررسی می‌کند و چنان که ذکر شد چنین سیستم‌هایی از رفتار منظم دور از آشوب، پایداری و استحکام بالا برخوردارند. بنابراین فاصله یک گراف علامت‌دار از تعادل ساختاری در تحلیل رفتار سیستم به عنوان یک شاخص ساختاری اهمیت دارد. معروف‌ترین معیاری که برای سنجش چنین فاصله‌ای مطرح شده است، شاخص ناخشنودی گراف است که تعریف آن در 1-2 ارائه شد. Iacono و همکاران این شاخص را برای انواعی از شبکه‌های علامت‌دار زیست‌مولکولی محاسبه نموده و نشان داده‌اند که اصولاً این شبکه‌ها ساختاری نزدیک به

تبادل دارند و در عین حال شبکه‌های تنظیم ژنی نسبت به شبکه‌های پیام‌رسانی سلولی ساختار از سطح تبادل ساختاری بالاتری برخوردارند [11]. در پژوهش حاضر بر این نکته تأکید می‌شود که با وجود سطح بالای تبادل ساختاری در شبکه‌های مورد بررسی، دو الگوی افراز بر مبنای چگالی و افراز بر مبنای تبادل ساختاری از یکدیگر فاصله دارند و این فاصله به عنوان یک ویژگی ساختاری در تحلیل شبکه‌های علامت‌دار پیشنهاد می‌شود. این ویژگی ساختاری را می‌توان با شاخص‌هایی مانند اختلاف متوسط چگالی خوشه‌های متبادل با متوسط چگالی انجمن‌ها سنجید. همچنین متوسط اختلاف بین نسبت پیوندهای منفی در انجمن‌ها و این نسبت در خوشه‌های متبادل را می‌توان به عنوان معیاری برای سنجش این ویژگی به کار برد.

در مقاله احسانی و همکاران تأثیر این ویژگی بر کارکردهایی نظیر انتشار اطلاعات و اغتشاش در سیستم بررسی شده و نشان داده شده است که اختلاف بین خوشه‌های متبادل و انجمن‌ها سبب تسهیل انتشار اطلاعات و همچنین پیشرفت اغتشاش در سیستم می‌شود [21]. برعکس، انطباق دو الگوی مزبور بر یکدیگر، درجه استحکام و پایداری بالاتری را نتیجه می‌دهد. به این ترتیب نتایج این بخش نشان می‌دهد که با وجود ساختار تقریباً متبادل شبکه‌های تنظیم ژنی، دو الگوی افراز بر مبنای چگالی و افراز بر مبنای تبادل ساختاری از یکدیگر کاملاً متفاوت بوده و این واقعیت ظرفیت بالای شبکه تنظیم ژنی در انتشار اطلاعات را نشان می‌دهد.

2-3- ارتباط بین خوشه‌های متبادل و نقش پیوندهای

مثبت

چنانکه پیش‌تر ذکر شد، خوشه‌بندی متبادل شبکه را به تعدادی خوشه یا به عبارتی به زیر واحدهای نزدیک به یکدیگر تجزیه می‌کند به طوری که بیشتر ارتباط‌های درون

خوشه‌ای از نوع مثبت و اکثریت ارتباط‌های بین خوشه‌ها از نوع منفی است. نقش مهم این ارتباطات منفی بین‌خوشه‌ای به عنوان پل در پایداری و استحکام سیستم در تعدادی از کارها بررسی شده است [12].

در این پژوهش نقش خاص ارتباطات مثبت واسطه بین این خوشه‌ها در حفظ تبادل سیستم بررسی شده است. در الگوریتم‌های افراز بر مبنای تبادل از این گروه اقلیت پیوندهای مثبت معمولاً به عنوان خطا یاد می‌شود. در مسیر تبادل سیستم پویا تحت قوانین ساده تعامل بین گره‌ها، وضعیت گره‌ها در جستجوی حداقل انرژی و کمینه کردن ناسازگاری، در جهت کم کردن تعداد یال‌های ناخشنود اصلاح می‌شود. به این ترتیب، گره‌های داخلی یک خوشه با ارتباطات غالباً مثبت به سمت یک وضعیت یکسان گرایش می‌یابند، در حالی که پیوندهای مثبت بین دو خوشه با وضعیت‌های متفاوت سبب انتقال وضعیت‌ها در جهت یکسان می‌شوند. چنین نقشی به عنوان واسطه در انتقال اطلاعات و انتشار اغتشاش در شبکه اهمیت می‌یابد. به این ترتیب می‌توان انتظار داشت که ایجاد اغتشاش در گره‌های این پیوندهای واسطه اثر بزرگتری در کل سیستم خواهد داشت. در این پژوهش این گزاره روی داده‌های تعدادی شبکه تنظیم ژنی مربوط به Dream 4 in silico challenge بررسی شده است.

این مجموعه داده از سایت GNW در آدرس gnw.sourceforge.net قابل دسترس است و شامل زیرشبکه‌های صدتایی و ده‌تایی استخراج شده از شبکه بازنویسی سلول‌های *E.coli* و *S.cerevisiae* می‌باشد. داده‌های سری زمانی سطح بیان ژنی مربوط به پنج شبکه صدتایی برای باینری کردن سطح بیان هر ژن استفاده شده است. سپس داده‌های مربوط به حالت تبادل و پاسخ سیستم به حذف تک تک ژن‌ها برای محاسبه فاصله حالت اولیه و نهایی به ازای حذف هر ژن استفاده شده است. در

محاسبه فاصله، دو معیار فاصله همینگ¹ و فاصله متعامد² مورد نظر بوده است. توضیحات مربوط به داده‌های مورد استفاده، الگوریتم باینری کردن و تعاریف فاصله‌ها در بخش مواد و روش‌ها آمده است.

به منظور بررسی تأثیر پیوندهای مثبت واسطه بین خوشه‌های متعادل بر سیستم، شبکه‌های یاد شده به خوشه‌های متعادل افراز و یال‌های مثبت واسطه شناسایی شد. بررسی و تحلیل پاسخ سیستم به حذف تک تک ژن‌ها نشان می‌دهد که حذف ژن‌های مربوط به این پیوندهای مثبت واسطه منجر به انحراف بزرگتری در شرایط سیستم نسبت به تعادل اولیه می‌شود. این نکته اهمیت این عناصر را در حالت کلی سیستم و کنترل مسیر تعادلی نشان می‌دهد. همچنین نتایج نشان می‌دهد که فاصله ایجاد شده بین حالت اولیه و نهایی به ازای حذف دو ژن متناظر پیوندهای مزبور تقریباً یکسان است و این مطلب تأییدی بر اهمیت پیوند است. خلاصه نتایج این قسمت در جدول 5 آمده است.

به این ترتیب عناصر مؤثر در کنترل سیستم را می‌توان از روی پیوندهای مثبت واسطه بین خوشه‌های متعادل شبکه شناسایی کرد. اجرای این روش روی شبکه net b. ژن‌هایی را که به طور مکرر در نتایج خوشه‌بندی متعادل به عنوان سر کمان‌های مثبت واسطه ظاهر می‌شوند را چنین تعیین می‌کند: *arcA* به عنوان تنظیم کننده پاسخ در یک سیستم تنظیم دوتایی، *crp* به عنوان فاکتور بازنویسی عمومی، واسطه مهارکننده کاتابولیت، *dcuR* به عنوان تنظیم کننده پاسخ در یک سیستم تنظیم دوتایی و *gadE* به عنوان فعال کننده بازنویسی.

4- خطوط پژوهشی آینده

چنانچه در این مقاله مطرح شد، در شبکه‌های علامت‌دار فاصله بین دو الگوی افراز بر مبنای چگالی و افراز بر

مبنای تعادل ساختاری را می‌توان به عنوان یک ویژگی ساختاری مؤثر در نوع کارکرد سیستم در نظر گرفت. ارزیابی انواع متفاوت شبکه‌های زیست‌مولکولی، همچنین ارزیابی شبکه‌های زیست‌مولکولی در فازهای متفاوت می‌تواند رابطه بین این شاخص ساختاری و رفتار سیستم در فازهای مختلف را آشکار کند. از کاربردهای متصور این مطلب در تحلیل و کنترل شبکه‌ها به عنوان مثال می‌توان به کنترل مسیر تمایز سلولی اشاره نمود.

به عنوان نکته دیگر، به کاربردهای متداول خوشه‌بندی ژن‌ها بر مبنای مشابهت تغییرات سطح بیان و کشف الگوهای مؤثر در کارکرد یا بیماری خاص اشاره می‌شود. به عنوان مثال الگوریتم‌های گوناگونی برای استخراج شبکه‌های همبستگی بر مبنای داده‌های بیان RNA اشاره کرد که خوشه‌ها را بر مبنای مشابهت نوسانات سطح بیان در یک فاصله زمانی تعیین می‌کند [22]. این رویکرد فقط وضعیت گره‌ها در طول زمان را در نظر می‌گیرد. به عنوان رویکرد جدید در آشکارسازی واحدهای کارکردی، خوشه‌بندی بر مبنای وضعیت گره‌ها و علامت پیوندها به صورت توأم پیشنهاد می‌شود که به تعیین گروه‌های نزدیک به تعادل ساختاری با گره‌های با وضعیت نوسانات مشابه در طول زمان می‌انجامد. این رویکرد در تعیین دقیق‌تر زیرسیستم‌های زیستی یا الگوهای تغییر در خوشه‌های ژنی مؤثر در بیماری‌های خاص می‌تواند مفید باشد.

5- مواد و روش‌ها

1-1-5 داده

1-1-5 شبکه بازنویسی سلول *E. coli*، داده‌های net a. و netb.
در این پژوهش از داده شبکه تنظیم بازنویسی سلول *E. coli* از سایت RegulonDb (release 8.5) در آدرس <http://regulondb.ccg.unam.mx/> استفاده شده است.

1. Hamming Distance
2. Orthogonal Distance

جدول 5 حذف ژن‌های مربوط به کمان‌های واسطه بین خوشه‌های متعادل سبب تغییرات بزرگ در پاسخ شبکه می‌شود.

فاصله همینگ	فاصله متعامد	سر کمان‌های واسطه	حداقل و حداکثر	متوسط فاصله	حداقل و حداکثر	متوسط فاصله	شماره
مربوط حذف	مربوط به حذف	بین خوشه‌ها	فاصله همینگ	همینگ	فاصله متعامد	متعامد	شبکه
این رئوس	این رئوس						
26, 18	13/8, 10/9	gadX, modE	2-31	8/3	3/9-16/0	5/65	زیر شبکه 1
28, 25	12/17, 8/53	fnr, lrp	10-28	16/64	4/9-12/17	6/35	زیر شبکه 2

این داده گراف علامت‌دار و جهت‌دار شبکه را فراهم می‌کند. این گراف 27 مؤلفه همبند ضعیف دارد و بزرگترین مؤلفه آن 1308 گره دارد که *net a* نامیده می‌شود.

اطلاعات عمومی *net a*: تعداد گره: 1308، تعداد کمان: 2600، تعداد کمان‌های منفی: 1237، تعداد کمان‌های مثبت: 1363، چگالی گراف: 0/0015، درجه متوسط گراف: 3.97

با حذف گره‌های با درجه خروجی برابر صفر گرافی با 155 گره بدست می‌آید. بزرگترین مؤلفه همبند ضعیف این گراف 85 گره دارد که *net b* نامیده می‌شود.

اطلاعات عمومی *net b*: تعداد گره: 85، تعداد کمان: 122، تعداد کمان‌های منفی: 48، تعداد کمان‌های مثبت: 122، چگالی گراف: 0/16، درجه متوسط گراف: 2/84، بیشترین درجه حداکثر: 34.

2-1-5- مجموعه داده Dream 4 in silico network challenge

این مجموعه داده از طریق وب سایت gnw.sourceforge.net در دسترس است [25-23] و شامل داده‌های مربوط به تعدادی زیرشبکه‌های 100 تایی و 10 تایی از شبکه تنظیم بازنویسی سلول *E.coli* و *S.cerevisiae* می‌شود. پویایی شبکه‌ها با استفاده از یک مدل دقیق شبکه تنظیم ژنی شبیه‌سازی شده است. فایل‌های *timeseries.tsv* پاسخ شبکه را به اغتشاشات مختلف و حذف این اغتشاشات در طول زمان نشان

می‌دهد. برای شبکه‌های اندازه 100 تایی، 10 سری زمانی که هریک شامل 21 نقطه زمانی است داده شده است. این داده برای باینری کردن سطح بیان ژن‌های مختلف استفاده شده است. فایل‌های *wildtype.tsv* سطوح بیان ژنی را در وضعیت پایا متناظر با نبود اغتشاش نشان می‌دهد. سطوح بیان در وضعیت پایا متناظر با حذف تک تک ژن‌ها در فایل‌های *knockouts.tsv* آورده شده است. همچنین سطوح بیان در وضعیت پایا متناظر با نصف کردن سطح بیان تک تک ژن‌ها در فایل‌های *knockdowns.tsv* داده شده است. در این پژوهش، این داده‌ها برای محاسبه فاصله حالت تعادل اولیه از حالت تعادل به ازای حذف تک تک ژن‌ها استفاده شده است.

3-1-5- الگوریتم باینری کردن داده‌های سطح بیان ژنی و روش محاسبه فاصله دو حالت شبکه

حالت شبکه (100 تایی) در هر نقطه زمانی با یک بردار 100 تایی نمایش داده می‌شود که درایه n ام آن برابر سطح بیان ژن n ام در آن لحظه زمانی است. به این ترتیب فاصله دو حالت شبکه (متناظر با دو وضعیت تعادلی) به عنوان فاصله دو بردار متناظر تعریف می‌شود. اگر مقادیر بردارها یعنی سطوح بیان به صورت پیوسته مقادیر بین صفر و یک باشد، فاصله دو بردار V_1 و V_2 به طور ساده با رابطه فاصله متعامد تعریف می‌شود:

$$d_{\text{Manhattan}} = \sum_i |v_1(i) - v_2(i)|$$

اگر سطوح بیان ژنی به صورت باینری متناظر حالت روشن و خاموش تعریف شود، فاصله دو بردار با فاصله

همینگ محاسبه می‌شود. فاصله همینگ دو بردار n تایی باینری برابر تعداد درایه‌های متناظر نامساوی در آن دو است.

بنابراین برای محاسبه فاصله همینگ، داده باید باینری شود. از آنجا که مدل‌های گسسته مانند مدل شبکه‌های بولین¹ از داده گسسته استفاده می‌کنند، در مورد نرمال‌سازی، گسسته‌سازی و باینری‌سازی داده‌های ژنی با توجه به ملاحظات خاص این نوع داده، پژوهش‌های متعددی وجود دارد. از آن جمله می‌توان به [28-26] اشاره کرد که مطالعات بعدی در زمینه باینری‌سازی داده ژنی بیشتر به الگوریتم ارائه شده در [28] ارجاع می‌دهند. در واقع این مرحله از اهمیت خاصی برخوردار است چراکه، گسسته‌سازی داده تا حدودی اطلاعات آن را به نحو غیر قابل بازگشتی کاهش می‌دهد. به همین دلیل تخصیص نادرست مقدار گسسته در مراحل بعدی سبب ناکارآمدی بسیاری از مدل‌ها و بسیاری روش‌های تحلیلی می‌شود. به عنوان مثال تخصیص نادقیق برای سیستمی که حساسیت بالا نسبت به شرایط اولیه یا نویز دارد، به عنوان شیفت در شرایط اولیه به شمار می‌آید. از طرفی دقت بسیار بالا مستلزم پیچیدگی محاسباتی است. بنابراین انتخاب روش نرمال‌سازی و باینری‌سازی بر حسب مدل و روش‌های بعدی باید دقت مناسبی داشته باشد.

به این منظور در این پژوهش از الگوریتم معرفی شده در [28] استفاده شده است و در آن اصلاحات جزئی صورت گرفته است. چنانکه در همان مقاله بیان شده است، از آنجا که دامنه سطح بیان ژن‌های مختلف با یکدیگر متفاوت است، انتخاب یک آستانه یکسان به عنوان مرز خاموش و روشن در ژن‌های متفاوت مناسب نیست. بنابراین الگوریتم مورد نظر پس از مرتب‌سازی صعودی سطوح بیان برای هر ژن به طور جداگانه، آستانه باینری را متناظر با اولین جهش در دنباله مرتب‌شده در نظر می‌گیرد. در

این پژوهش با اصلاح مختصری به جای اولین جهش، بزرگترین جهش به عنوان آستانه جداسازی خاموش و روشن در نظر گرفته شده است. شبه کد الگوریتم مورد نظر در اینجا داده شده است. بردار g متناظر مقادیر مختلف سطح بیان یک ژن خاص در فواصل زمانی است. خروجی الگوریتم با بردار b نمایش داده شده است و مقادیر باینری متناظر سطح بیان ژن را در این فواصل زمانی نشان می‌دهد.

```
Gene expression vector:  $g$ , length( $g$ ):  $n$ 
 $s \leftarrow \text{sort}(g)$ 
for  $i=1:n-1$ 
 $d(i) \leftarrow s(i+1) - s(i)$ 
end for
 $j \leftarrow \text{index}(\max(d))$ 
for  $i=1:n$ 
if  $g(i) \geq s(j)$  then
 $b(i) \leftarrow 1$ 
else
 $b(i) \leftarrow 0$ 
end if
end for
```

5-2- نرم‌افزار

کد نویسی برنامه‌های این پژوهش با نرم‌افزار MATLAB 2010 انجام شده است. برای تحلیل شبکه‌ها و افراز گراف از نرم‌افزار Pajek 3.13 استفاده شده است. برای تعیین انجمن‌ها از الگوریتم Louvain در این نرم‌افزار در حالت multi-level coarsening- multi level refinement تنظیم پارامتر resolution برابر $0/3$ و $0/7$ استفاده شده است. برای تعیین خوشه‌های متعادل هم از الگوریتم Doreian- Mrvar در این نرم‌افزار استفاده شده است.

6- منابع

- [1] A.L. Barabasi, Z.N. Oltvai, (2004) "Network biology: Understanding the cell's functional organization", Nature Reviews Genetics, Vol. 5, 101-114
- [2] U. Sauer, M. Heinemann, N. Zamboni, (2007)"GENETICS: Getting Closer to the Whole Picture". Science, 316 (5824): 550-551.
- [3] Snoep, Jacky L; Westerhoff, Hans V. (2005). "From isolation to integration, a systems

- [18] Singh, R., Dasgupta, S., & Sinha, S. (2014). Extreme variability in convergence to structural balance in frustrated dynamical systems. *EPL (Europhysics Letters)*, 105(1), 10003.
- [19] V.D. Blondel, J.-L. Guillaume, R. Lambiotte and E. Lefebvre (2008). "Fast unfolding of community hierarchies in large networks". *J. Stat. Mech.* (10): P10008.
- [20] Doreian, P., Mrvar, A. (1996). A partitioning approach to structural balance. *Social networks*, 18(2), 149-168.
- [21] Ehsani, M., Sepehri, M.M. (2014). Balanced clusters and diffusion process in signed networks, *Journal of Industrial and Systems Engineering*, Vol.7, Issue 1, 104-117.
- [22] Doreian, P., & Mrvar, A. (2009). Partitioning signed social networks. *Social Networks*, 31(1), 1-11.
- [23] Marbach D, Schaffter T, Mattiussi C. and Floreano D (2009) Generating Realistic *In Silico* Gene Networks for Performance Assessment of Reverse Engineering Methods. *Journal of Computational Biology*, 16(2):229–239.
- [24] Stolovitzky G, Prill RJ, Califano A (2009) Lessons from the DREAM2 Challenges. In Stolovitzky G, Kahlem P, Califano A, Eds, *Annals of the New York Academy of Sciences*, 1158:159–95.
- [25] Stolovitzky G, Monroe D, Califano A (2007) Dialogue on Reverse- Engineering Assessment and Methods: The DREAM of High-Throughput Pathway Inference. In Stolovitzky G and Califano A, Eds, *Annals of the New York Academy of Sciences*, 1115:11–22.
- [26] Zhou X., Wang X., and Dougherty E.D. (2003) Binarization of Microarray Data on the Basis of a Mixture Model, Vol. 2, 679–684, July Molecular Cancer Therapeutics 679.
- [27] Irizarry R.A., Hobbs B., Collin F., Speed T.P., (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data, *Biostatistics*, 4, 2, pp. 249–264
- [28] Shmulevich, I., & Zhang, W. (2002). Binary analysis and optimization-based normalization of gene expression data. *Bioinformatics*, 18(4), 555-565.
- [29] Butte, A. J., & Kohane, I. S. (2000). Mutual information relevance networks: functional genomic clustering using pairwise entropy measurements. In *Pac Symp Biocomput*, Vol. 5, pp. 418-429.
- biology approach for building the Silicon Cell". In Alberghina, Lilia; Westerhoff, Hans V. *Systems Biology: Definitions and Perspectives*. Topics in Current Genetics 13. Berlin: Springer-Verlag. pp. 13–30.
- [4] Alon, U. (2007). Network motifs: theory and experimental approaches. *Nature Reviews Genetics*, 8(6), 450-461.
- [5] Shen-Orr, S. S., Milo, R., Mangan, S., & Alon, U. (2002). Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nature genetics*, 31(1), 64-68.
- [6] Girvan, M., & Newman, M. E. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12), 7821-7826.
- [7] Smith, H. L. (2008). *Monotone dynamical systems: an introduction to the theory of competitive and cooperative systems* (Vol. 41). American Mathematical Soc..
- [8] Sontag ED. (2007). Monotone and near-monotone biochemical networks. *Systems and Synthetic Biology*, 1:59-87.
- [9] Cartwright, D., & Harary, F. (1956). Structural balance: a generalization of Heider's theory. *Psychological review*, 63(5), 277
- [10] Davis, J. A. (1967). Clustering and structural balance in graphs. *Human relations*. Vol 20(2), 181-187
- [11] Iacono, G., & Altafini, C. (2010). Monotonicity, frustration, and ordered response: an analysis of the energy landscape of perturbed large-scale biological networks. *BMC systems biology*, 4(1), 83.
- [12] Lin, C. C., Lee, C. H., Fuh, C. S., Juan, H. F., & Huang, H. C. (2013). Link clustering reveals structural characteristics and biological contexts in signed molecular networks. *PloS one*, 8(6), e67089.
- [13] Heider, F. (1946). Attitudes and cognitive organization. *Journal of Psychology* 21, 107–112.
- [14] Antal T., Krapivsky P.L., and Redner S. (2005) "Dynamics of social balance on networks", *Physical Review E*, 72 Issue 3
- [15] Traag, V. A., Van Dooren, P., & De Leenheer, P. (2013). Dynamical models explaining social balance and evolution of cooperation. *PloS one*, 8(4), e60063.
- [16] Marvel, S. A., Strogatz, S. H., & Kleinberg, J. M. (2009). Energy landscape of social balance. *Physical review letters*, 103(19), 198701.
- [17] Altafini, C. (2012). Dynamics of opinion forming in structurally balanced social networks. *PloS one*, 7(6), e38135.