دو فصلنامه محاسبات و سامانه‌های توزیع‌شده
www.ijdcs.ir

تعیین برهم‌کنش‌های پایدار بر اساس ژن‌های خانه‌دار

مسلم محمدی جنقرا*[1]، کوثر عباس خواه[1]
[1]دانشگاه پیام نور، تهران

چکیده

ژن‌های خانه‌دار در تمامی شرایط زیستی و مراحل رشد بافت حاضر بوده و همیشه بیان می‌شوند. ژن خانه‌دار، ژن بنیادی است که برای نگهداری از عملکرد اولیه سلول موردنیاز بوده و در بیشتر سلول‌های یک ارگانیسم بیان‌شده و پروتئین‌های متناظر را می‌سازد. در حالت عادی، برخی از ژن‌ها در شرایط و گام‌های متفاوت از چرخه سلولی بیان نمی‌شوند که می‌تواند دینامیک حضور یک پروتئین و به‌تبع آن برهمکنش‌های گذرا یا پایدار را مشخص کند. برهمکنش‌های پایدار در تعیین کمپلکس‌های پروتئینی نقش مهمی دارند. یکی از چالش‌های موجود در زیست‌شناسی سیستمی، تعیین حد آستانه‌های مناسب برای تعیین پروتئین‌های فعال در هر زمان و مشخص کردن برهم‌کنش‌های پایدار هست. هدف ما در این مقاله، تعیین حد آستانه منحصر به هر ژن برای مشخص کردن ژن‌های خانه‌دار و برهم‌کنش‌های پایدار هست. با استفاده از الگوریتم بهینه‌سازی فرا ابتکاری کرم شب‌تاب و تابع جذابیت مبتنی بر ترکیب مجموعه کمپلکس‌های استاندارد و بیان هم‌زمان ژن‌ها، حد آستانه مخصوص هر ژن تعیین می‌شود. نتایج تجربی روی‌داده‌های موجود نشان می‌دهد که حد آستانه‌های ایجادشده با روش ابداعی نتایج بهتری نسبت به روش‌های قبلی داشته است.

کلمات کلیدی: برهم‌کنش پروتئین-پروتئین، برهم‌کنش پایدار، الگوریتم کرم شب‌تاب، ژن‌های خانه‌دار، آستانه گذاری

Email: Mo_mohammadi@pnu.ac.ir

تاریخچه مقاله:
تاریخ ارسال: ۹۸/۱۰/۱۵
تاریخ اصلاحات: ۹۸/۱۱/۲۰
تاریخ پذیرش: ۹۸/۱۲/۱۵
تاریخ انتشار: ۹۸/۱۲/۲۰

Stable Interactions Detection Based on Housekeeping Genes

Moslem Mohammadi Jenghara*[1], Kosar Abbaskhah[1]
[1]Payame Noor University, Tehran, Iran

Abstract

Housekeeping genes are expressed in all biological conditions and stages of tissue growth. A housekeeping gene is the basic gene required for maintaining the cell's primary function and is expressed in most cells of an organism to produces corresponding proteins. Normally, some genes are not expressed in different conditions and steps of the cell cycle that can determine the dynamics of the presence of a protein and consequently transient or stable interactions. Stable interactions play an important role in the determination of protein complexes. One of the challenges in system biology is to determine the appropriate thresholds to determine active proteins at each time point and to identify stable interactions. Our goal in this paper is to determine the unique threshold for each gene to characterize housekeeping genes and stable interactions. Determination of the specific threshold for each gene is determined by using the firefly optimization algorithm with the attraction function based on the combination of the standard protein complex and co-expression of genes. Experimental results on the existing dataset show that the thresholds created by the proposed method had better results than the previous methods

Keywords:
Protein-protein interaction
Stable interactions
Firefly algorithm
Housekeeping gene
Thresholding

روش ارجاع به مقاله:
تعیین برهم‌کنش‌های پایدار بر اساس ژن‌های خانه‌دار، م.محمدی جنقرا، ک. عباس خواه، دو فصلنامه محاسبات و سامانه‌های توزیع‌شده، سال دوم، شماره دوم، شماره پیاپی ۴ ، سال ۱۳۹۸، ص ۵۶-۶۵

۱- مقدمه

بیان ژن[1] فرآیند ساخته‌شدن محصولات ژنی، عمدتاً پروتئین، از اطلاعات ژن هست. بیان ژن روند رونویسی[2] توالی DNA به‌توالیmRNA هست که بعداً به‌توالی اسیدآمینه به نام پروتئین ترجمه می‌شود. تعداد نسخه‌های تولیدشده از RNA سطح بیان ژن نامیده می‌شود. هر ژن طبیعی دارای یک نرخ از سطح بیان هست. به‌طور دقیق برخی از ژن‌ها فاقد کدهای تولیدکننده‌ی پروتئین هستند که در شرایط و گام‌های متفاوت از چرخه سلولی می‌تواند دینامیک حضور یک پروتئین را مشخص کند. زمانی که دو یا بیش از دو پروتئین به یکدیگر متصل شوند برهم‌کنش پروتئین–پروتئین رخ می‌دهد، که غالباً عملکردهای بیولوژیکی را انجام می‌دهند. بسیاری از فرایندهای مولکولی مهم در سلول مانند تکثیر DNA توسط کمپلکس‌های بزرگ مولکولی انجام می‌شوند. این کمپلکس‌ها از تعداد زیادی پروتئین‌های سازمان‌دهی شده به‌وسیله برهم‌کنش بین پروتئین‌ها ساخته‌شده‌اند. برهم‌کنش-های بین پروتئین‌ها برای اکثر عملکردهای زیست‌شناسی مهم هستند. به‌عنوان‌مثال، سیگنال‌های سطح خارجی یک سلول توسط برهم‌کنش پروتئینی بین مولکول‌های سیگنالی به داخل سلول هدایت می-شوند [۱]

برهم‌کنش‌های پروتئینی با توجه به دوره حیاتشان می‌توانند به دو گروه پایدار و گذرا تقسیم شوند. برهم‌کنش‌های پایدار در برازندگی ماندگاری سلول‌ها مهم بوده و پایداری آن‌ها دائمی و برگشت‌ناپذیر هست. درحالی‌که برهم‌کنش‌های گذرا وابسته به زمان و شرایط هستند که مکانیسمی برای پاسخ سریع سلول به تحریکات خارجی را فراهم می‌کنند [۲]. بررسی کمپلکس‌های پروتئینی زمانی می‌تواند بعد جدیدی به مکانیسم پیمانه‌ای پویا بازکرده و درک ما را از دلایل بیماری‌ها بهبود بخشد. هرچند کمپلکس‌های زمانی مختلف در نقاط زمانی متفاوتی رخ می‌دهند، ولی کمپلکس‌های پروتئینی زیادی وجود دارند که ماکرو مولکول‌های پایداری در راستای انجام عملکرد زیستی مهم را تشکیل می‌دهند. برهم‌کنش‌های پایدار زیادی که دارای نقش اساسی برای سلول هستند، در نقاط زمانی مختلف حفظ‌شده و کمپلکس‌های متناظر آن‌ها نیز در شبکه‌های پویای برهم‌کنش پروتئینی پشت سر هم ظاهر می‌شوند. به خاطر حفظ برازندگی و پایداری سلول و همچنین برای اجتناب از اختلال نامطلوب در عملکرد اساسی سلول، این کمپلکس‌ها باید دارای تغییرات صاف و ملایمی در طول زمان باشند [۳]. برهم‌کنش‌های پایدار را می‌توان با استفاده از ژن‌های خانه‌دار استخراج کرد.

تنوع شگفت‌انگیز موجودات زنده ناشی از الگوهای مختلف بیان ژن‌ها در بافت‌های مختلف است. اگرچه بیشتر ژن‌ها در زیرمجموعه‌ای از بافت‌ها بیان می‌شوند، اما برخی از محصولات ژن برای حفظ عملکرد سلول پایه لازم است و به‌طور اساسی در کلیه سلول‌های موجودات یافت می‌شوند. ژن‌های خانه‌دار در تمامی شرایط زیستی و مراحل رشد بافت حاضر بوده و همیشه بیان می‌شوند. در زیست‌شناسی مولکولی، ژن خانه‌داری به‌طورمعمول ژن بنیادی است که برای نگهداری از عملکرد اولیه سلول موردنیاز بوده و در تمام سلول‌های یک ارگانیسم در شرایط عادی و

[1] Gene expression
[2] Transcription

فیزیولوژیکی بیان‌شده است. کمبود و یا جهش غیر مترادف ژن خانه‌داری به‌احتمال‌زیاد به بیماری منجر خواهد شد. به‌طورمعمول، ژن خانه‌داری به‌عنوان کنترل‌های مولکولی در اندازه‌گیری‌های کیفی یا نیمه کمی بیان ژن وفق داده‌شده است [۴]. در این مقاله رهیافتی برای تعیین ژن‌های خانه‌دار با استفاده از تجمیع داده‌های بیان ژن و کمپلکس‌های پروتئینی ارائه‌شده است. با استفاده از ژن‌های خانه‌دار، برهم‌کنش‌های پایدار بین پروتئینی استخراج‌شده است.

ریزآرایه‌ها غالباً برای شناسایی مجموعه‌هایی از ژن‌ها که به‌صورت همه‌جا یا در بافت یا شرایط خاص بیان می‌شوند، مورداستفاده قرار می‌گیرند. بااین‌حال، این تکنیک ازلحاظ فنی مستقل نبوده و مستعد تأثیرپذیری است [۶]. در اصل، شناسایی مجموعه ژن‌های خانه‌دار با استفاده از داده‌های ریزآرایه ساده است. فقط باید ژن‌هایی را جستجو کرد که در همه بافت‌ها و شرایط آزمایشگاهی بیان‌شده باشند. بااین‌حال، اگرچه ژن‌های خانه‌داری به‌طور یکنواخت در انواع خاصی از سلول بیان می‌شوند، اما می‌توانند در سایرین متفاوت باشند، بنابراین، مطالعه و انتخاب ژن‌های کنترل مناسب برای نمونه‌های بالینی بیمار برای تجزیه‌وتحلیل بیان ژن بسیار مهم است [۵].

ازآنجاکه آزمایش نوردرن بلات برای هر ژن در هر بافت غیرعملی است، برای اعتبارسنجی هر لیست از ژن‌های خانه‌دار، آزمایش مستقل لازم است. ایسنبرگ و همکارانش [۶]، یک روش اعتبار سنجی برای ژن‌های خانه‌دار ارائه کرده‌اند که از یک حد آستانه تعیین‌شده برای ژن‌های بابیان بالا استفاده می‌کند. فراری و همکاران [۷]، روش مبتنی بر رده‌بندهای بیز ساده، برای کاوش ژن‌های خانه‌دار بر اساس خصوصیات فیزیکی و عملکردی ژن‌ها مانند طول اگزون و اندازه‌گیری فشردگی کروماتین ارائه کرده‌اند. در تحقیقات پانینا و همکارانش [۸] برای ارزیابی پایداری ژن‌های موردمطالعه در تمام شرایط تجربی و مناسب بودن این ژن‌ها به‌عنوان مرجع برای آزمایش qPCR، از پنج الگوریتم آماری بر روی‌داده‌های بیان ژن استفاده‌شده است. هررا و همکارانش از حد آستانه گذاری برای اعتبارسنجی ژن‌های خانه‌دار به‌منظور ارزیابی بیان ژن piscirickettsia salmonis استفاده کرده‌اند [۹].

۲- تعیین پروتئین‌های فعال در نقاط زمانی مختلف

داده‌های بیان ژن معمولاً برای تعیین حضور یک پروتئین به‌وسیله تطبیق دادن با یک حد آستانه استفاده می‌شوند. سطوح دینامیک بیان ژن، دینامیک حضور پروتئین را مشخص می‌کند ولی بیان تضمینی برای دینامیک برهم‌کنش نمی‌دهد. به‌عبارت‌دیگر، هرچند دو پروتئین ممکن است در یک‌زمان حضورداشته باشند ولی هیچ تضمینی نیست که حتماً در همان زمان باهم تعامل داشته‌اند. به این دلیل که ممکن است در آن زمان یکی از پروتئین‌ها در حالت غیرفعال بوده و نمی‌توانسته فعالیتی انجام دهد و با دیگران در تعامل باشد. اگر سطح بیان ژن کمتر از یک حد آستانه‌ای باشد، به‌منزله عدم حضور این پروتئین هست؛ اما بعضی از پروتئین‌های مهم وجود دارند که مقدار کمی از سطح بیان دارند پس بنابراین نمی‌توان فقط یک حد آستانه برای همه پروتئین‌ها برای تعیین حضور یا عدم حضور و فعالیت یا عدم فعالیت در نظر

گرفت [۱]؛ بنابراین برای هرکدام از ژن‌ها باید حد آستانه مجزا تعیین شوند.

هر پروتئینی که مقدار سطح بیان ژن آن از حد آستانه بیشتر باشد، به‌عنوان پروتئین فعال در شبکه حضور خواهد داشت؛ و این کار برای تمامی نقاط زمانی نمونه‌برداری انجام خواهد شد؛ یعنی اگر کمترین مقدار بیان ژن، بزرگ‌تر از حد آستانه باشد، آن ژن همیشه فعال خواهد بود. در سوی دیگر، اگر بزرگ‌ترین مقدار بیان ژن از حد آستانه کمتر باشد بنابراین آن ژن غیرفعال خواهد بود [۱۲].

با تجمیع داده‌های کمپلکس‌های پروتئینی و داده‌های بیان ژن در نقاط زمانی مختلف حد آستانه‌هایی برای تعیین پروتئین‌های فعال استخراج می‌شوند.

۲-۱- روش‌های مرسوم 3σ و 0.7

داده‌های بیان ژن برای نقاط زمانی $(1 \le t \ge T)$ t، سنجیده می‌شوند. برهم‌کنش‌های بین پروتئینی بر اساس فعال بودن هم‌زمان دو پروتئین مرتبط ساخته می‌شود [۱]. در نقطه زمانی t، پروتئینی فعال است اگر مقدار بیان ژن آن از یک حد آستانه تعریف‌شده‌ای (AT(i)) بیشتر باشد. حد آستانه به‌صورت زیر تعیین می‌شود.

$$AT(i) = u(i) + 3\sigma(i)(1 - F(i)) \qquad \text{(فرمول-۱)}$$

u(i)بیانگر میانگین بیان ژن پروتئین i ام هست.

$$F(i) = 1 / (1 + \sigma^2(i)) \qquad \text{(فرمول-۲)}$$

F(i) تابع وزن هست که منعکس‌کننده نوسانات بیان پروتئین i ام هست. روش فوق تحت عنوان 3σ شناخته می‌شود. تانگ و همکارانش [۱۱] برای تعیین حد آستانه‌بالا از مقدار ثابت 0.7 استفاده کرده‌اند.

۳- روش پیشنهادی

۳-۱- تعیین حد آستانه با استفاده از الگوریتم کرم شب‌تاب (FA_thr)

در روش‌های موجود، از حد آستانه‌های ثابت برای همه ژن‌ها استفاده‌شده است، ولی هدف ما در این مقاله تعیین حد آستانه منحصر به هر ژن هستیم. در این بخش از مقاله با استفاده از اطلاعات اضافی موجود، همچون پروفایل‌های بیان ژن در زمان‌ها و شرایط مختلف و مجموعه کمپلکس‌های پروتئینی استاندارد طلایی سعی در تعیین حد آستانه‌های مناسب شده است. برای انجام کار، مسئله را به یک مسئله بهینه‌سازی تبدیل کرده و با استفاده از الگوریتم بهینه‌سازی فرا ابتکاری کرم شب‌تاب آن را حل کرده‌ایم. از مهم‌ترین بخش‌های کار، تعیین تابع جذابیت در الگوریتم کرم شب‌تاب است که در اینجا تابعی از ترکیب مجموعه کمپلکس‌های استاندارد و بیان هم‌زمان ژن‌ها تعریف‌شده است. مراحل کلی روش پیشنهادی (FA_thr) در الگوریتم (۱) آورده شده است. بعد از استخراج حد آستانه‌های مجزا برای هر ژن و اعمال آن بر روی مجموعه داده‌های بیان ژن، ژن‌های فعال و ژن‌هایی که در تمامی نقاط زمانی و در تمامی شرایط سلولی بیان‌شده‌اند، تعیین می‌شوند. این ژن‌ها تحت عنوان ژن‌های خانه‌دار شناخته‌شده و در تعیین برهم‌کنش‌های پایدار استفاده‌شده‌اند. جزئیات پارامتریک روش پیشنهادی در کار پیشین ما [۱۳] در دسترس هست.

(الگوریتم-۱): مراحل استخراج حد آستانه برای هر ژن، مبتنی بر الگوریتم کرم شب‌تاب

Algorithm FA_thr: Gene expression threshold determination based on Firefly Algorithm
Input: PPI network (PPIN),
Gene expression data (GSE),
Gold standard protein complexes.
Output: Gene Expression Thresholds,
Permanent protein-protein interactions.
Cox ← Compute gene Co-expression matrix from GSE data,
//coxi expression matrices are created using the bicor criterion [12].
FCOX ← Integrating gene Co-expression matrices (Cox),
// weighting each cox_i using information of gold standard protein complexes, such as CYC2008
Seed ←select Seed genes from high expressed genes at different time points.
Thr ← Create randomly thresholds for each gene,
While (Termination condition not satisfied)
Attractiveness ← Compute attractiveness based on seed genes and Thr.
FA_thr ← Apply Firefly algorithm to optimize the random Thr.
Candidate Housekeeping genes← Apply FA_thr thresholds on the gene expression data,
Determine permanent PPI based on Candidate Housekeeping genes and Biogrid dataset,

۴- نتایج تجربی

۴-۱- مجموعه داده‌ها

BioGRID [۱۵]: این پایگاه داده، مجموعه‌ای یکپارچه و پیوسته در حال به‌روزرسانی از برهم‌کنش‌های فیزیکی و کلی هست. در این مقاله از نسخه (3.2.121) برای ارگانیسم Saccharomyces_cerevisiae استفاده‌شده است. این مجموعه بیش از ۳۴۵۰۰۰ برهم‌کنش داشته و بیش از ۲۷ ارگانیسم مختلف را شامل می‌شود. این مجموعه با داشتن ۲۳۰۰۰۰ برهم‌کنش غیرتکراری از مخمر بزرگ‌ترین مجموعه PPI برای این ارگانیسم هست. جدول (۱) جزئیات مربوط با این مجموعه داده را نشان می‌دهد.

GEO [۱۶]: سطوح بیان ژن در نقطه‌های زمانی مختلف اندازه‌گیری شده و در مجموعه داده‌هایی ذخیره می‌شود. پایگاه داده GEO این سری‌های مختلف بیان ژن را تحت پلتفرم‌های مشخص (GPLxxx) و برای نمونه‌های مختلفی (GSMxxx) تحت اسامی یکتایی (GSExxx) ذخیره می‌کند. برای مثال: سری GSE3431 از پلتفرم GPL90 در ۱۲ نقطه زمانی برای و فواصل زمانی ۲۵ دقیقه‌ای اندازه‌گیری شده است. ۶۷۷۷ سطح بیان ژن اندازه‌گیری شده است. داده‌های PPI استفاده‌شده در این پژوهش از مجموعه داده‌های biogrid هست. همچنین اطلاعات مربوط به بیان ژن و بیان هم‌زمان نیز از GEO برداشته‌شده است. جدول (۲) بعضی ازسری‌های بیان ژن را نشان می‌دهد.

(جدول-۱): جزئیات اطلاعات برهم‌کنش و تعداد پروتئین‌های ارگانیسم Saccharomyces cerevisiae در مجموعه داده Biogrid

Organism	Saccharomyces cerevisiae		
Experiment Type	PHYSICAL	GENETIC	COMBINED
Raw Interactions	141200	208332	349532
Non-Redundant Interactions	86656	150424	230333
Unique Proteins	6442	5679	6657
Unique Publications	7601	7392	12661

(جدول-۲): لیست منتخبی ازسری‌های بیان ژن

Series name	Number of series
Gse26169	210
Gse25582	151
Gse18121	42
Gse15254	72
Gse11452	170
Gse9482	40
Gse7645	48
Gse3431	36
Gse3076	96

مجموعه استاندارد CYC2008 [14] شامل ۴۰۸ کمپلکس پروتئینی با پوشش ۱۶۲۷ پروتئین هست. این مجموعه شامل ۱۴۹ کمپلکس با اندازه بزرگ‌تر از ۴ پروتئین هست. جدول (۳) اطلاعات آماری این مجموعه کمپلکس‌های استاندارد طلایی را نشان داده است. (جدول-۳): اطلاعات آماری مجموعه داده CYC2008

Number of proteins in complex (#P)	Number of complexes containing #P proteins
2	172
3	87
4	44
5	21
6	21
7	10
8	12
>8	41

برای ارزیابی روش پیشنهادی از میزان تغییرات پروتئین‌ها در نقاط زمانی مختلف بیان ژن‌ها و همچنین مجموعه برهم‌کنش‌های مجموعه داده Biogrid استفاده‌شده است. میزان تغییرات برهم‌کنش‌ها و حضور پروتئین‌ها، نسبت به نقاط زمانی قبل از خود، هم از جنبه تغییرات حضور پروتئینی و هم ازلحاظ وجود برهم‌کنش بین آن‌ها اندازه‌گیری شده است. نتایج در جدول (۴) و جدول (۵) آورده شده است.

همان‌طوری که در جدول (۴) مشخص است نرخ تغییرات در دو روش پویاسازی 0.7 و 3σ خیلی بالا هست؛ و حتی در برخی نقاط زمانی، نزدیک به یک می‌شود. برای مثال میزان تغییرات نقطه زمانی ۸ نسبت به نقطه زمانی ۷، در روش‌های 0.7 و 3σ و روش پیشنهادی به ترتیب 0.97، 0.91 و 0.19 هست. نرخ تغییر 0.97 بیانگر این است که تقریباً تمامی پروتئین‌های فعال موجود در یک نقطه زمانی، جدید بوده و از مرحله قبل هیچ پروتئینی وجود ندارد و این به معنی تغییرات کلی در سلول هست که با مفاهیم زیست‌شناسی در تناقض هست. به خاطر حفظ برازندگی و پایداری سلول و همچنین برای اجتناب از اختلال نامطلوب در عملکرد اساسی سلول، این کمپلکس‌ها باید دارای تغییرات صاف و ملایمی در طول زمان باشند. نتایج روش پیشنهادی، تغییرات صاف و ملایمی را نشان می‌دهد.

(جدول-۴): تغییرات حضور پروتئین‌ها در دنباله نقاط زمانی (میزان تغییرات در نقطه زمانی t نسبت به t-1)

time	3sigma	0.7	FA_thr
1	0.74	0.90	0.11
2	0.52	0.46	0.20
3	0.50	0.47	0.09
4	0.48	0.39	0.07
5	0.45	0.29	0.10
6	0.42	0.53	0.12
7	0.46	0.50	0.10
8	0.91	0.97	0.19
9	0.42	0.76	0.15
10	0.65	0.54	0.05
11	0.43	0.51	0.07
12	0.64	0.88	0.12
13	0.71	0.82	0.06
14	0.50	0.51	0.18
15	0.49	0.37	0.09
16	0.74	0.66	0.22
17	0.65	0.76	0.12
18	0.72	0.86	0.09
19	0.84	0.97	0.18
20	0.65	0.91	0.12
21	0.44	0.61	0.05
22	0.58	0.47	0.03
23	0.68	0.55	0.08
24	0.78	0.73	0.12
25	0.67	0.78	0.10
26	0.49	0.55	0.20
27	0.47	0.54	0.07
28	0.34	0.59	0.13
29	0.52	0.25	0.08
30	0.81	0.87	0.13
31	0.90	0.98	0.16
32	0.60	0.90	0.15
33	0.52	0.59	0.05
34	0.30	0.34	0.05
35	0.41	0.36	0.04
average	0.58	0.63	0.11

(جدول-۵): تغییرات حضور برهم‌کنش پروتئین‌ها در دنباله نقاط زمانی (میزان تغییرات در نقطه زمانی t نسبت به t-1)

time	3sigma	0.7	FA_thr
1	0.95	0.99	0.39
2	0.76	0.78	0.52
3	0.85	0.77	0.36
4	0.72	0.65	0.34
5	0.74	0.59	0.39
6	0.74	0.80	0.43
7	0.71	0.74	0.35
8	0.99	1.00	0.58
9	0.73	0.96	0.41
10	0.94	0.83	0.33
11	0.74	0.85	0.32
12	0.81	0.93	0.33
13	0.91	0.96	0.34
14	0.74	0.80	0.48
15	0.85	0.73	0.37
16	0.93	0.90	0.69
17	0.83	0.93	0.40
18	0.94	0.95	0.39
19	0.98	1.00	0.53
20	0.89	0.98	0.38
21	0.65	0.89	0.30
22	0.88	0.66	0.27
23	0.93	0.83	0.36
24	0.88	0.79	0.34
25	0.91	0.96	0.35
26	0.75	0.85	0.50
27	0.75	0.80	0.35
28	0.67	0.85	0.41
29	0.80	0.60	0.35
30	0.94	0.98	0.41
31	0.99	1.00	0.53
32	0.86	0.99	0.40
33	0.83	0.92	0.30
34	0.66	0.56	0.31
35	0.72	0.56	0.29
average	0.83	0.84	0.39

جدول (۵)، نرخ تغییرات برهم‌کنش‌های بین پروتئین های فعال را نشان می‌دهد. همان‌طور که در جدول مشخص است، تغییرات روش پیشنهادی نسبت به روش‌های قبلی صاف‌تر هست. از نکات مثبت و جالب روش پیشنهادی، تعیین برهم‌کنش‌های پایدار در طول نقاط زمانی مختلف به‌طور ضمنی هست؛ یعنی در

حین تعیین مقدار حد آستانه، این مقدار برای برخی از پروتئین‌ها برابر صفر و یا یک عدد نزدیک به صفر تعیین می‌شد، به نحوی که تمامی بیان ژن‌ها در نقاط زمانی از این مقدار بزرگ‌تر بوده و این پروتئین‌ها همیشه فعال بودند. برهم‌کنش بین آن‌ها به‌عنوان برهم‌کنش‌های پایدار مشخص می‌شوند.

۵- نتیجه‌گیری

در این مقاله، تعیین حد آستانه‌های مناسب برای تعیین پروتئین‌های فعال و همچنین تعیین ژن‌های خانه‌دار به‌عنوان یکی از چالش‌های موجود در زیست‌شناسی سیستمی، مورد بررسی واقع‌شده و روش جدیدی برای تعیین حد آستانه ارائه شد. یکی از نکات مهم در روش پیشنهادی تعیین حد آستانه منحصر به هر ژن هست؛ و همه حد آستانه‌ها از فرمول و رابطه ثابتی برای همه ژن‌ها استفاده نمی‌کنند. همچنین ادغام روش بهینه‌سازی فرا ابتکاری الگوریتم کرم شب‌تاب با مفاهیم بیان هم‌زمان ژن‌ها و مجموعه داده‌های استاندارد طلایی از دیگر نکات بارز کار هست.

۶- مراجع

[1] *J. Wang, X. Peng, M. Li, Y. Luo, and Y. Pan, "Active protein interaction network and its application on protein complex detection," in Bioinformatics and Biomedicine (BIBM), IEEE International Conference on, 2011, pp. 37-42: 2011.*

[2] *L. Ou-Yang, DQ. Dai, XL. Li, M. Wu, XF. Zhang and P. Yang, "Detecting temporal protein complexes from dynamic protein-protein interaction networks". J. BMC bioinformatics, vol. 15, no. 1, p, 335, 2014.*

[3] *S. Byrum, SK. Smart, S. Larson and AJ. Tackett, "Analysis of stable and transient protein–protein interactions", J. Chromatin Remodeling. Humana Press. Springer. pp. 143-152, 2012.*

[4] *Pan, J. B. Hu, S. C. Shi, D. Cai, M. C. Li, Y. B. Zou, Q. & Ji, Z. L. PaGenBase: a pattern gene "database for the global and dynamic understanding of gene function". PloS one, 8(12), e80747, 2013.*

[5] *Lee, S. R. Jo, M. J. Lee, J. E. Koh, S. S. & Kim, S. Y. "Identification of novel universal housekeeping genes by statistical analysis of microarray data". BMB Reports, 40(2), 226-231, 2007.*

[6] *Eisenberg, E. & Levanon, E. Y, "Human housekeeping genes are compact". TRENDS in Genetics, 19(7), 362-365, 2003.*

[7] *De Ferrari, L. & Aitken, S. "Mining housekeeping genes with a Naive Bayes classifier". BMC genomics, 7(1), 277, 2006.*

[8] *Panina, Y. Germond, A. Masui, S. & Watanabe, T. M. "Validation of common housekeeping genes as reference for qpcr gene expression analysis during iPS reprogramming process". Scientific reports, 8(1), 8716, 2018.*

[9] *Flores-Herrera, P. Arredondo-Zelada, O. Marshall, S. H. & Gómez, F. A. "Selection and validation of reliable housekeeping genes to evaluate Piscirickettsia salmonis gene expression". Infection, Genetics and Evolution, 63, 151-157, 2018.*

[10] *Q. Xiao, J. Wang, X. Peng, and F.X. Wu, "Detecting protein complexes from active protein interaction networks constructed with dynamic gene expression profiles," Proteome science, vol. 11, no. supplement 1, p. S20, 2013.*

[11] *X. Tang, J. Wang, B. Liu, M. Li, G. Chen, and Y. Pan, "A comparison of the functional modules identified from time course and static PPI network data,"*

BMC bioinformatics, vol. 12, no. 1, p. 339, 2011.

[12] *L. Song, P. Langfelder, and S. Horvath, "Comparison of co-expression measures: mutual information, correlation, and model based indices," BMC bioinformatics, vol. 13, no. 1, p. 328, 2012.*

[13] *Jenghara, M. M. Ebrahimpour-Komleh, H. & Parvin, H. "Dynamic protein–protein interaction networks construction using firefly algorithm". Pattern Analysis and Applications, 21(4), 1067-1081, 2018.*

[14] *L. Ou-Yang, D.Q. Dai, X.L. Li, M. Wu, X.F. Zhang, and P. Yang, "Detecting temporal protein complexes from dynamic protein-protein interaction networks," BMC bioinformatics, vol. 15, no. 1, p. 335, 2014.*

[15] *A. Chatr-Aryamontri et al. "The BioGRID interaction database: 2015 update," Nucleic acids research, vol. 43, no. D1, pp. D470-D478, 2014.*

[16] *R. Edgar, M. Domrachev, and A. E. Lash, "Gene Expression Omnibus: NCBI gene expression and hybridization array data repository," Nucleic acids research, vol. 30, no. 1, pp. 207-210, 2002.*

مسلم محمدی جنقرا مدرک کارشناسی خود را در رشته مهندسی کامپیوتر در سال ۱۳۸۴ از دانشگاه آزاد، مدرک کارشناسی ارشد خود را در رشته مهندسی کامپیوتر گرایش هوش مصنوعی در سال ۱۳۸۷ از دانشگاه علم و صنعت و مدرک دکتری مهندسی کامپیوتر گرایش هوش مصنوعی را در سال ۱۳۹۷ از دانشگاه کاشان دریافت کرده است. ایشان در حال حاضر به عنوان هیات علمی دانشگاه پیام نور مشغول به کار می باشد. زمینه های پژوهشی مورد علاقه ایشان عبارت اند از: داده کاوی، بیوانفورماتیک، پردازش متن، شبکه های پیچیده

کوثر عباس خواه مدرک کارشناسی را در رشته ی علوم تربیتی از دانشگاه فرهنگیان اخذ کرده است و در حال حاضر دانشجوی کارشناسی ارشد رشته آموزش و بهسازی منابع انسانی بوده و در آموزش و پرورش شاغل می باشد.