



دانشگاه صنعتی امیرکبیر

(پلی تکنیک تهران)

دانشکده مهندسی کامپیوتر و فناوری اطلاعات

پایان نامه کارشناسی ارشد

گرایش معماری کامپیوتر

عنوان

تجزیه‌ی سلسله‌مراتبی شبکه‌های متابولیکی جهت محاسبه‌ی

مدهای پایه شار

نگارش

سینا قادرمرزی

استاد راهنما

دکتر مرتضی صاحب‌الزمانی

استاد مشاور

دکتر سید امیر مرعشی

بهمن ۱۳۹۵

اینجانب سینا قادرمرزی متعهد می‌شوم که مطالب مندرج در این پایان نامه حاصل کار پژوهشی اینجانب تحت نظارت و راهنمایی اساتید دانشگاه صنعتی امیرکبیر بوده و به دستاوردهای دیگران که در این پژوهش از آنها استفاده شده است مطابق مقررات و روال متعارف ارجاع و در فهرست منابع و مآخذ ذکر گردیده است. این پایان نامه قبلاً برای احراز هیچ مدرک هم سطح یا بالاتر ارائه نگردیده است.

در صورت اثبات تخلف در هر زمان، مدرک تحصیلی صادر شده توسط دانشگاه از درجه اعتبار ساقط بوده و دانشگاه حق پیگیری قانونی خواهد داشت.

کلیه نتایج و حقوق حاصل از این پایان نامه متعلق به دانشگاه صنعتی امیرکبیر می‌باشد. هرگونه استفاده از نتایج علمی و عملی، واگذاری اطلاعات به دیگران یا چاپ و تکثیر، نسخه برداری، ترجمه و اقتباس از این پایان نامه بدون موافقت کتبی دانشگاه صنعتی امیرکبیر ممنوع است. نقل مطالب با ذکر مآخذ بلامانع است.

امضا

## قدردانی

در این جا لازم است از استاد گرانقدر دکتر صاحب الزمانی، که در همه‌ی مراحل این تحقیق از حمایت‌ها و زحمات بی‌دریغ ایشان بهره برده‌ام، قدردانی نمایم. بی‌شک دست‌آوردهای این پایان نامه مرهون راهنمایی‌ها و کمک‌های صبورانه‌ی ایشان است. همچنین از استاد گرامی جناب آقای دکتر مرعشی، که راهنمایی‌های ارزشمندشان نقش به‌سزایی در هدایت این تحقیق به مسیر مناسب داشته است، نهایت تشکر را دارم.

## تقدیم به

پدر و مادر عزیزم که در همه‌ی مراحل زندگی به من دلگرمی دادند.

## چکیده

مدهای پایه‌ی شار به عنوان تعریف ریاضی مفهوم مسیره‌های متابولیکی و نیز به عنوان ابزارهایی در مهندسی متابولیک، دارای اهمیت فراوانی هستند. به همین دلیل تحقیقات فراوانی بر روی محاسبه‌ی کارآمد آن‌ها صورت می‌گیرد. اما با روش‌های محاسباتی موجود، محاسبه‌ی آن‌ها برای شبکه‌های با بیش از چند صد واکنش، حتی بر روی بسترهای پردازشی بزرگ غیر عملی است. ایده‌ی تجزیه‌ی این شبکه‌ها به زیرشبکه‌های کوچکتر توسط تحقیقات متعددی دنبال شده است. اما این تحقیقات زیرشبکه‌ها را به تنهایی مورد بررسی قرار می‌دهند و روشی برای اتصال مدهای پایه‌ی زیرشبکه‌ها ارائه نکرده‌اند. در تحقیق حاضر سعی شده است با استفاده از تعریف درشت‌واکنش‌ها، ارتباط بین مدهای پایه‌ی زیرشبکه‌ها و مدهای پایه‌ی شبکه‌ی کلی برقرار شود. در این روش مساله‌ی نحوه‌ی اتصال مدهای پایه‌ی زیرشبکه‌ها به تشکیل یک شبکه‌ی درشت‌دانه می‌انجامد. با محاسبه‌ی مدهای پایه‌ی شبکه‌ی درشت‌دانه شده می‌توان مدهای پایه‌ی شبکه‌ی اصلی را یافت. بنابراین محاسبه‌ی مدهای پایه‌ی یک شبکه‌ی بزرگ، به محاسبه‌ی مدهای پایه‌ی یک شبکه‌ی کوچکتر کاهش می‌یابد. روش تجزیه و درشت‌دانه‌سازی شبکه به این شکل، می‌تواند به عنوان یک روش فشرده‌سازی شبکه به کار رود. نتایج اعمال این روش بر روی چهار شبکه‌ی متابولیکی واقعی با اندازه‌های مختلف نشان داده‌اند که این روش، اندازه‌ی شبکه را بیش از ۵۰ درصد کاهش می‌دهد. همچنین در مورد شبکه‌های متابولیکی بزرگ، اندازه‌ی ماتریس کرنل، که یکی از شاخص‌های اصلی زمان محاسبه‌ی مدهای پایه است، بین ۳۰ تا ۴۰ درصد کاهش یافته است. به علاوه، برای یک شبکه‌ی با اندازه‌ی متوسط، زمان محاسبه‌ی مدهای پایه به نصف کاهش یافته است؛ که نشان می‌دهد که استفاده از روش درشت‌دانه‌سازی ارائه شده می‌تواند به کاهش زمان اجرا منجر گردد. ضمن این که در همه‌ی موارد، روش درشت‌دانه‌سازی منجر به کاهش بیشتری در تعداد متابولیت‌ها و واکنش‌ها نسبت به روش‌های موجود گردیده است.

## واژه‌های کلیدی:

شبکه‌های متابولیکی، فشرده‌سازی شبکه، تجزیه‌ی شبکه، مد پایه‌ی شار، درشت‌واکنش

|  |    |
|--|----|
| ۱ فصل اول مقدمه.....   | ۱  |
| ۲ فصل دوم مفاهیم پایه.....   | ۶  |
| ۱-۲ متابولیسم و مدل های آن.....  | ۷  |
| ۱-۱-۲ مدل مبتنی بر معادلات دیفرانسیل.....  | ۸  |
| ۲-۱-۲ مدل های مبتنی بر محدودیت.....  | ۸  |
| ۳-۱-۲ مدل مبتنی بر گراف.....   | ۹  |
| ۲-۲ مدل مبتنی بر محدودیت و تحلیل های مهم آن.....                                       | ۱۰ |
| ۱-۲-۲ واکنش ها و شار آن ها.....  | ۱۰ |
| ۲-۲-۲ ماتریس استوکیومتری.....  | ۱۱ |
| ۳-۲-۲ قیدهای حالت پایا.....  | ۱۱ |
| ۴-۲-۲ واکنش های تبادلی و متابولیت های خارجی.....                                       | ۱۲ |
| ۵-۲-۲ واکنش های برگشت ناپذیر.....  | ۱۴ |
| ۶-۲-۲ فضای جواب و مخروط شار.....   | ۱۴ |
| ۷-۲-۲ تحلیل های مربوط به مدل مبتنی بر محدودیت.....                                     | ۱۵ |
| ۳-۲ مدهای پایه ی شار.....  | ۱۸ |
| ۴-۲ روش توصیف دوتایی برای محاسبه ی مدهای پایه شار.....                                 | ۱۹ |
| ۳ فصل سوم پیشینه پژوهش.....  | ۲۲ |
| ۱-۳ مدهای پایه ی شار و محاسبه.....   | ۲۳ |
| ۱-۱-۳ روش های مبتنی بر هندسه ی محاسباتی.....   | ۲۳ |
| ۲-۱-۳ ابزارهای اولیه و بهبودهای ایجاد شده.....   | ۲۳ |
| ۳-۱-۳ استفاده از روش های مبتنی بر گراف.....  | ۲۵ |
| ۴-۱-۳ پیچیدگی زمانی مسائل مربوط به مدهای پایه.....                                     | ۲۶ |
| ۵-۱-۳ محاسبه ی زیرمجموعه ای از مدهای پایه.....   | ۲۷ |
| ۶-۱-۳ محاسبه ی مدهای پایه ی مولد برای یک زیرمجموعه ی خاص از فضای شار.....              | ۳۰ |
| ۲-۳ تجزیه ی شبکه های متابولیکی.....  | ۳۲ |
| ۱-۲-۳ بررسی ساختار اتصالات در شبکه های متابولیکی.....                                  | ۳۳ |
| ۲-۲-۳ تجزیه ی شبکه های متابولیکی به منظور یافتن ماژول های عملیاتی.....                 | ۳۳ |
| ۳-۲-۳ تجزیه ی شبکه های متابولیکی و مدهای پایه.....                                     | ۳۵ |
| ۴ فصل چهارم روش پیشنهادی.....  | ۳۷ |
| ۱-۴ به دست آوردن چارچوبی برای تجزیه ی شبکه های متابولیکی به زیرشبکه های جدا از هم..... | ۴۰ |
| ۱-۱-۴ نحوه ی تعریف زیرشبکه ها.....   | ۴۰ |

|    |   |
|----|---|
| ۴۳ | ۲-۱-۴ نحوه‌ی تجزیه‌ی شبکه.....  |
| ۴۷ | ۲-۴ به دست آوردن مدهای پایه از طریق درشت‌دانه سازی.....                           |
| ۴۹ | ۱-۲-۴ تبدیل مدهای پایه به واکنش‌های متناظر.....                                   |
| ۴۹ | ۲-۲-۴ ارتباط بین یک شبکه با شبکه‌ای متشکل از درشت‌واکنش‌های آن.....               |
| ۵۳ | ۳-۲-۴ درشت‌دانه‌سازی شبکه.....  |
| ۵۵ | ۴-۲-۴ مدهای پایه‌ی شبکه‌ی درشت‌دانه.....  |
| ۵۵ | ۵-۲-۴ نسبت مسیرهای به دست آمده با مدهای پایه.....                                 |
| ۵۸ | ۳-۴ به وجود آوردن شرط جدایی زیرشبکه‌ها با افزودن گره‌های ساختگی.....              |
| ۶۰ | ۴-۴ مسیرهای موازی.....  |
| ۶۲ | ۵-۴ روند کلی روش تجزیه به ازای یک افراز دلخواه از شبکه و محاسبه‌ی مدهای پایه..... |
| ۶۴ | ۶-۴ نحوه‌ی انتخاب زیرشبکه‌ها.....   |
| ۶۷ | <b>۵ فصل پنجم نتایج و تحلیل آن‌ها.....</b>  |
| ۶۸ | ۱-۵ شرایط آزمایش.....   |
| ۶۹ | ۲-۵ نتایج اجرا.....   |
| ۷۶ | ۳-۵ تحلیل نتایج.....  |
| ۷۹ | <b>۶ فصل ششم جمع‌بندی و کارهای آینده.....</b>                                     |
| ۸۲ | <b>منابع و مراجع.....</b>   |

## صفحه

## فهرست اشکال

|  |    |
|--|----|
| شکل ۱-۲ یک واکنش و نحوه‌ی نمایش آن.....  | ۱۰ |
| شکل ۲-۲ یک شبکه‌ی متابولیکی و ماتریس استوکیومتری آن.....                               | ۱۱ |
| شکل ۳-۲ شبکه‌ی متابولیکی با اضافه کردن واکنش‌های تبدیلی.....                           | ۱۳ |
| شکل ۴-۲ شبکه‌ی متابولیکی با اضافه کردن متابولیت‌های خارجی به همراه ماتریس مربوطه.....  | ۱۳ |
| شکل ۵-۲ فضای جواب حاصل از معادلات و نامعادلات حاصل (مخروط شار) [۱].....                | ۱۵ |
| شکل ۱-۴ ایده‌ی معادل‌سازی یک مسیر متابولیکی با یک واکنش.....                           | ۳۹ |
| شکل ۲-۴ یک شبکه‌ی متابولیکی نمونه.....   | ۴۱ |
| شکل ۳-۴ مدهای پایه برای زیرشبکه‌ی متناظر با مجموعه‌ی متابولیت‌های 6، 7، 8 و 9.....     | ۴۲ |
| شکل ۴-۴ تجزیه‌ی به دست آمده از افراز متابولیت‌ها برای شبکه‌ی شکل ۲-۴.....              | ۴۷ |
| شکل ۵-۴ شبکه‌ی متشکل از درشت‌واکنش‌ها برای زیرشبکه‌ی $N_3$ در شکل ۴-۴.....             | ۵۰ |
| شکل ۶-۴ یک توزیع شار نمونه که به دو شکل قابل تجزیه است.....                            | ۵۱ |
| شکل ۷-۴ مدهای پایه‌ی شبکه‌ی شکل ۶-۴.....   | ۵۲ |
| شکل ۸-۴ شبکه‌ی درشت‌دانه متناظر با تجزیه‌ی موجود در شکل ۴-۴.....                       | ۵۴ |
| شکل ۹-۴ یک زیرشبکه‌ی مشخص شده به همراه مدهای پایه‌ی آن.....                            | ۵۷ |
| شکل ۱۰-۴ شبکه‌ی درشت‌دانه برای شبکه‌ی نشان‌داده شده در شکل ۹-۴.....                    | ۵۸ |
| شکل ۱۱-۴ نحوه‌ی افزودن متابولیت‌های ساختگی برای هر یک از واکنش‌های بین زیرشبکه‌ای..... | ۵۹ |



## صفحه

## فهرست جداول

|  |    |
|--|----|
| جدول ۱-۵ نتایج آزمایش ۱.....   | ۷۰ |
| جدول ۲-۵ نتایج آزمایش ۲.....   | ۷۲ |
| جدول ۳-۵ نتایج آزمایش ۳.....   | ۷۴ |
| جدول ۴-۵ نتایج آزمایش ۴.....   | ۷۵ |
| جدول ۵-۵ مقایسه‌ی نتایج روش‌های فشرده‌سازی موجود با روش درشت‌دانه‌سازی شبکه..... | ۷۶ |

۱

## فصل اول

### مقدمه

## مقدمه

زیست‌شناسی، علم مطالعه موجودات زنده، در حال ورود به عصر جدیدی است که به کمک پیشرفت‌های علوم کامپیوتر و فناوری اطلاعات، همانند علوم فیزیک و شیمی تبدیل به یک علم کمی گردد. امکاناتی که عصر حاضر از لحاظ آسانی پردازش و دسترسی به اطلاعات فراهم کرده است و همچنین پیشرفت‌های نظری علوم کامپیوتر موجب به وجود آمدن افق‌های جدیدی پیش روی این علم گشته‌اند. زمینه‌هایی مانند زیست‌شناسی محاسباتی<sup>۱</sup> و بیوانفورماتیک<sup>۲</sup> با استفاده از این امکانات، علی‌رغم پیچیدگی‌های فراوان سیستم‌های زیستی، موفق شده‌اند برای پدیده‌های مختلف زیستی مدل‌های کمی ارائه نمایند. نتیجه‌ی چنین دستاوردهایی تحول در فرایندهایی از قبیل درمان بیماری‌ها و کشف داروها می‌باشد. البته یکی از مهم‌ترین پیامدهای مهم این دستاوردها گسترده شدن هرچه بیشتر دانش بشری نسبت به جهان زنده از طریق شتاب یافتن اکتشافات جدید در عرصه‌ی علم زیست‌شناسی است. چنین افق‌هایی باعث شده‌اند که اهمیت این تحقیقات روز به روز بیشتر شود.

متابولیسم<sup>۳</sup> سیستمی از واکنش‌های شیمیایی است که از طریق آن‌ها فعالیت‌های حیاتی سلول از جمله سوخت و ساز صورت می‌گیرد. این سیستم اهمیت ویژه‌ای در سازوکارهای مختلف موجود زنده دارد به‌طوری که بخش قابل توجهی از زیست‌شناسی محاسباتی به بررسی متابولیسم می‌پردازد. متابولیسم به روش‌های مختلفی مدل می‌شود که هریک از این مدل‌ها میزان متفاوتی از جزئیات را شامل می‌شود. یکی از مهمترین مدل‌هایی که در مورد متابولیسم به کار گرفته می‌شود مدل مبتنی بر محدودیت<sup>۴</sup> است. این مدل به بررسی حالت‌های پایایی<sup>۵</sup> ممکن برای شار واکنش‌های یک شبکه می‌پردازد. در این نوع مدل با اضافه کردن فرض حالت پایا، بررسی متابولیسم به یک سری معادلات و نامعادلات جبری خطی بین شار واکنش‌ها ساده شده است. همچنین این روش از لحاظ میزان جزئیات و قدرت پیش‌بینی نسبت به

---

<sup>1</sup> Computational Biology

<sup>2</sup> Bioinformatics

<sup>3</sup> Metabolism

<sup>4</sup> Constraint Based Model

<sup>5</sup> Steady State

مدل‌های دیگر متابولیسم رویکرد میانه‌ای را در پیش گرفته است. با توجه به این که سلول زنده اکثر اوقات در حالت پایا قرار دارد، فرض حالت پایا فرض محدودکننده‌ای نبوده و این مدل از لحاظ قابلیت پیش‌بینی موفق عمل نموده است. در نتیجه بیشترین تحقیقات در زمینه‌ی متابولیسم در چارچوب این مدل انجام می‌پذیرد.

مجموعه معادلات و نامعادلات جبری که در مدل مبتنی بر محدودیت به وجود می‌آیند عموماً به یک فضای جواب (و نه یک جواب منحصر به فرد) می‌انجامند. در واقع این فضای جواب، حالت‌های مجاز شار واکنش‌ها را در حالت پایا معرفی می‌نماید، بدون این که به معرفی یک وضعیت متابولیکی خاص بپردازد. برای بررسی و تحلیل این مجموعه جواب و استفاده از اطلاعاتی که این مدل ارائه می‌دهد، تحلیل‌های مختلفی وجود دارد. بخش قابل توجهی از این تحلیل‌ها سعی دارند که مجموعه جواب را به مسیرهای پایه تجزیه نمایند. ایده‌ی این رویکرد این است که متابولیسم را نباید به صورت یکپارچه بلکه باید به صورت مجموعه مسیرهای متابولیکی بررسی کرد. این روش‌ها به طور کلی به تحلیل مسیرهای متابولیکی<sup>۶</sup> معروفند. روش تحلیل مسیر با استفاده از مدهای پایه‌ی شار، اصلی‌ترین ابزار در این زمینه است. مدهای پایه‌ی شار<sup>۷</sup> مسیرهای حداقلی و غیر قابل تجزیه از واکنش‌ها هستند که می‌توانند در حالت پایا فعال باشند.

ایده‌ی مطالعه‌ی متابولیسم با استفاده از مسیرهای متابولیکی<sup>۸</sup> ایده‌ی جدیدی نیست و متابولیسم از مدت‌ها پیش به همین صورت مطالعه شده است، به طوری که زیست‌شناسان و بیوشیمی‌دان‌ها با استفاده از شناخت خود مسیرهای ساخت و تجزیه‌ی مواد معروف را مشخص کرده‌اند. در واقع مدهای پایه‌ی شار هم‌اکنون به عنوان تعریف دقیق ریاضی برای مفهوم سنتی مسیر متابولیکی پذیرفته شده‌اند و این بر اهمیت آن‌ها می‌افزاید. اهمیت مفهومی و کاربردی مدهای پایه‌ی شار باعث شده است که تحقیقات زیادی برای تولید ابزارهای مناسب برای محاسبه‌ی آن‌ها انجام گیرد.

<sup>۶</sup> Metabolic Pathway Analysis

<sup>۷</sup> Elementary Flux Modes (EFMs)

<sup>۸</sup> Metabolic Pathways

با وجود تلاش‌های زیاد برای محاسبه‌ی سریع و کارآمد محاسبه‌ی مدهای پایه‌ی شار، هم اکنون استفاده از آن‌ها محدود به شبکه‌های با اندازه‌ی کوچک می‌باشد (شبکه‌هایی با حداکثر یکصد واکنش). دلیل این موضوع نیاز بالای روش‌های محاسبه‌ی موجود به سرعت و حافظه و همچنین تعداد بسیار زیاد این مسیرها در شبکه‌های بزرگ می‌باشد.

از طرفی بررسی متابولیسم به صورت جامع معمولاً نیاز به بررسی شبکه‌های بزرگ دارد. زیست‌شناسی سیستم‌ها<sup>۹</sup>، به عنوان رویکردی نسبتاً جدید در زیست‌شناسی محاسباتی، بر مبنای این واقعیت شکل گرفته است که برای شناخت یک سیستم، شناخت اجزای آن کافی نمی‌باشد. در عوض باید بر هم کنش بین این اجزا را نیز در نظر گرفت و آن را به صورت مجتمع مورد بررسی قرار داد. بر همین مبنای زیست‌شناسی سیستم‌ها عمدتاً با شبکه‌های بسیار بزرگ، که مکانیسم‌ها را در سطوح بالا، همانند سلول، ارگانیسم و حتی جامعه بررسی می‌کنند سرو کار دارد. ساخت مدل‌های متابولیسم مقیاس ژنوم<sup>۱۰</sup> را، که در دهه‌های اخیر انجام گرفته است، می‌توان در راستای همین تفکر دانست. این مدل‌ها از روی اطلاعات ژنوم یک موجود زنده نقشه‌ی کاملی از متابولیسم را در یک مدل گردآوری می‌کنند. این شبکه‌های بزرگ عموماً شامل بیش از هزار واکنش هستند. بنابراین روش‌های موجود به هیچ عنوان قادر به محاسبه‌ی مدهای پایه‌ی شار در این شبکه‌ها نیستند.

چالش‌های استفاده از روش تحلیل مسیرهای متابولیکی و به طور خاص تحلیل مدهای پایه‌ی شار، در شبکه‌های مقیاس ژنوم، موضوعی است که در تحقیقات فراوانی به آن پرداخته می‌شود و نیاز به ایده‌ها و روش‌هایی برای پردازش این شبکه‌های بزرگ هم‌اکنون یک نیاز روز می‌باشد. در همین راستا موضوع این پژوهش نیز محاسبه‌ی مدهای پایه‌ی شار در شبکه‌های بزرگ است.

شبکه‌های بزرگ در زمینه‌های دیگر علمی نیز مورد توجه هستند. در همه‌ی زمینه‌های مهندسی کامپیوتر و به طور ویژه در مورد طراحی مدارهای پرتراکم دیجیتال<sup>۱۱</sup>، ایده‌ی طراحی ماژولار رویکرد طبیعی است که در مورد سیستم‌های پیچیده و بزرگ مورد استفاده قرار می‌گیرد. این کار به این صورت

<sup>9</sup> Systems Biology

<sup>10</sup> Genome-Scale Metabolic Models

<sup>11</sup> Digital VLSI Design

انجام می‌گیرد که همواره سعی می‌شود یک سیستم پیچیده را از ترکیب سیستم‌های کوچکتر به وجود آورند. در فرایند طراحی مدارهای پرتراکم دیجیتال به جای طراحی فیزیکی یک مدار به صورت یک پارچه، این کار به صورت سلسله‌مراتبی روی بخش‌های مختلف مدار انجام می‌شود. بدون چنین رویکردی پیاده‌سازی مدارهایی با پیچیدگی کنونی غیرممکن است. بررسی ماژولار متابولیسم نیز ایده‌ی جدیدی نیست و دانشمندان به طور سنتی متابولیسم را با استفاده از شهود به بخش‌هایی تقسیم کرده‌اند.

هدف این پژوهش ارائه‌ی یک روش محاسبه‌ی مدهای پایه‌ی شار در شبکه‌های متابولیکی بر مبنای ایده‌ی تجزیه‌ی شبکه است. این کار به این منظور انجام می‌شود که با بهره‌برداری از ساختار طبیعی ماژولار سلسله‌مراتبی<sup>۱۲</sup> در شبکه‌های متابولیکی، بتوان محاسبه‌ی مدهای پایه‌ی شار را تسهیل نمود. رویکرد مورد نظر این است که با استفاده از تجزیه‌ی شبکه‌ی متابولیکی، محاسبه‌ی مدهای پایه‌ی زیرشبکه‌ها و استفاده از آن‌ها برای ساختن مدهای پایه‌ی شبکه‌ی کلی، در مجموع عملیات ساده‌تر گردد. در فصل دوم پایان‌نامه، مفاهیم موردنیاز برای فصل‌های بعدی پایان‌نامه ارائه می‌شوند. در فصل سوم کارهای گذشته مرتبط با این پژوهش مورد بررسی قرار می‌گیرند. در فصل چهارم، روش پیشنهادی با جزئیات توضیح داده شده است. در فصل پنجم نتایج اعمال روش بر روی شبکه‌های نمونه را ارائه می‌دهیم. و در فصل ششم تحلیل نتایج و جمع‌بندی ارائه خواهد شد.

---

<sup>12</sup> Hierarchical Modularity

۲

فصل دوم

مفاهیم پایه

## مفاهیم پایه

در این قسمت تعاریف و مفاهیمی که برای درک بخش‌های بعدی پایان نامه لازم است، آورده شده است.

### ۱-۲ متابولیسم و مدل‌های آن

متابولیسم به مجموعه‌ای از واکنش‌های شیمیایی که در سلول‌های زنده با دخالت آنزیم‌های اختصاصی در حال وقوع هستند اطلاق می‌گردد. آنزیم‌ها موادی هستند که وجود آن‌ها برای وقوع خود به خودی واکنش‌های زیستی لازم است اما در این واکنش‌ها مصرف نمی‌شوند. بنابراین از طریق آنزیم‌ها یک سلول کنترل کاملی بر روی واکنش‌های پیش‌رونده در آن را دارد. این واکنش‌های زیستی در کنار یکدیگر شبکه‌ای پیچیده را تشکیل می‌دهند که از طریق آن‌ها مکانیسم‌های حیاتی سلول زنده، مانند سوخت و ساز، رشد و تکثیر انجام می‌گیرد. با توجه به تعداد زیاد و پیچیدگی ارتباطات این واکنش‌ها، متابولیسم به طور سنتی به صورت مسیرهای متابولیکی مطالعه شده است. یعنی مسیریایی از واکنش‌ها که طی آن‌ها یک فرایند چند مرحله‌ای شیمیایی به وقوع می‌پیوندد. این مسیرها البته از یکدیگر مستقل نیستند و امروزه با توجه به اهمیت این واقعیت، متابولیسم را در قالب شبکه‌های متابولیکی مطالعه می‌کنند.

یک شبکه‌ی متابولیکی از مجموعه‌ای ماده (اصطلاحاً متابولیت<sup>۱۳</sup>) و واکنش‌های بین آن‌ها تشکیل شده است. هر واکنش شامل تبدیل یک یا تعدادی متابولیت به متابولیت (های) دیگر است که با نسبت‌های کاملاً معین انجام می‌گیرد. این نسبت‌های معین اصطلاحاً نسبت‌های استوکیومتری نامیده می‌شوند. سرعت این تبدیل را شار<sup>۱۴</sup> واکنش می‌گوییم. معمولاً مساله‌ی اصلی مورد مطالعه در متابولیسم بررسی شار واکنش‌ها و نیز غلظت متابولیت‌ها است، هرچند جنبه‌های ساختاری و توپولوژیکی این شبکه‌ها نیز مورد مطالعه قرار می‌گیرند.

<sup>13</sup> Metabolite

<sup>14</sup> Flux



به طور کلی می‌توان مدل‌های متابولیسم را به دو دسته‌ی پویا<sup>۱۵</sup> و ساختاری<sup>۱۶</sup> تقسیم کرد. در مدل‌های پویا هدف بررسی تغییرات متابولیسم در طول زمان است در حالی که در مدل‌های ساختاری زمان مطرح نمی‌باشد. اغلب مطالعات بر روی متابولیسم در قالب مدل‌های زیر صورت می‌گیرد.

- مدل‌های مبتنی بر معادلات دیفرانسیل
- مدل‌های مبتنی بر محدودیت<sup>۱۷</sup> (حالت پایا)
- مدل‌های مبتنی بر گراف

در ادامه توضیح مختصری درباره‌ی هریک از این مدل‌ها ارائه می‌شود.

## ۲-۱-۱ مدل مبتنی بر معادلات دیفرانسیل

در این نوع مدل‌ها هدف بررسی دقیق شار واکنش‌ها و غلظت واکنش‌گرها در طول زمان است. این نوع مدل‌ها بیشتر به منظور بررسی نحوه‌ی پاسخ سلول به تغییرات ناگهانی در سیستم می‌پردازند. از آنجایی که ساختن چنین مدل‌هایی نیازمند به اندازه‌گیری دقیق پارامترها با استفاده از اندازه‌گیری‌های تجربی است، مدل‌های در دسترس از این نوع بسیار محدود هستند. زیرا اندازه‌گیری این پارامترها در بسیاری از موارد مشکل است.

## ۲-۱-۲ مدل‌های مبتنی بر محدودیت

این نوع مدل به بررسی فضای حالات مجاز متغیرها در حالت پایا می‌پردازد. یعنی بنابر روابط استوکیومتری واکنش‌ها (یا محدودیت‌های ذاتی دیگر مانند برگشت‌پذیری واکنش‌ها) حالت‌های پایای ممکن برای سیستم بررسی می‌شود. شرط حالت پایا به معنی آن است که غلظت مواد ثابت بوده و هر متابولیت با همان سرعتی که تولید می‌شود مصرف گردد. متابولیسم یک سلول در اکثر مواقع در یک

<sup>15</sup> Dynamic

<sup>16</sup> Structural

<sup>17</sup> Constraint Based Model

حالت پایا به سر می‌برد. بنابراین برای بررسی قابلیت‌های متابولیکی سلول، چنین فرضی می‌تواند اساس کار قرار گیرد.

معمولاً تنها اطلاعات لازم برای ساخت مدل مبتنی بر محدودیت، واکنش‌ها و ضرایب استوکیومتری و نیز برگشت‌پذیری هریک از واکنش‌ها است. این امر باعث شده که مدل‌های فراوانی با قابلیت اطمینان زیاد از این نوع ساخته شوند. ضمن اینکه مدل‌های از این نوع، اخیراً با استفاده از اطلاعات ژنتیکی موجودات زنده‌ی مختلف طی یک فرایند نیمه خودکار ساخته شده‌اند [۲، ۳]. مدل‌های مقیاس ژنوم متابولیسم مدل‌های بسیار بزرگی از این نوع هستند که معمولاً شامل کلیه‌ی واکنش‌ها و متابولیت‌های ممکن در یک موجود زنده بوده و برای موجودات مختلفی از جمله انسان موجود و در دسترس هستند [۴]. ظهور چنین مدل‌هایی امکان بررسی سراسری و جامع متابولیسم را ممکن ساخته‌اند و تحقیقات بسیاری بر روی این مدل‌ها در جریان است.

## ۲-۱-۳ مدل مبتنی بر گراف

در مدل‌های مبتنی بر گراف بر اساس ارتباطات متابولیت‌ها یا واکنش‌ها گراف‌های ساده، گراف‌های جهت دار، ابرگراف<sup>۱۸</sup> یا ابرگراف جهت‌دار ساخته می‌شوند [۵، ۶]. در این نوع مدل‌ها ساختار ارتباطات، معمولاً به منظور مطالعه‌ی جنبه‌های تکاملی یا ماژولار بودن، مورد بررسی قرار می‌گیرد. همچنین این نوع مدل‌ها برای مسیریابی جهت تخمین مسیرهای متابولیکی نیز به کار می‌روند [۷].

بیشتر مطالعات در زمینه‌ی متابولیسم بر روی مدل مبتنی بر محدودیت انجام می‌گیرد. مدهای پایه‌ی شار به عنوان مسیرهای حداقلی از واکنش‌ها که در حالت پایا می‌توانند فعالیت کنند. در قالب همین مدل تعریف می‌شوند. بنابراین روش مورد بررسی در این پژوهش نیز می‌تواند در همین قالب دسته بندی گردد. در ادامه جزئیات بیشتر این مدل ارائه می‌گردد.

<sup>18</sup> Hypergraph

## ۲-۲ مدل مبتنی بر محدودیت و تحلیل‌های مهم آن

مدل مبتنی بر محدودیت در حال حاضر یکی از پرکاربردترین مدل‌های متابولیسم است که به بررسی حالات ممکن برای شار واکنش‌ها در حالت پایا می‌پردازد. اطلاعات مورد نیاز برای ساخت این مدل‌ها معمولاً تنها شامل ضرایب استوکیومتری واکنش‌ها و نیز تعیین برگشت‌ناپذیر بودن هریک از واکنش‌ها است. این مدل شامل تعدادی متابولیت و تعدادی واکنش است که متغیرهای مساله شار واکنش‌ها هستند.

### ۲-۲-۱ واکنش‌ها و شار آن‌ها

هر واکنش متابولیکی شامل تبدیل یک یا چند ماده به یک یا چند ماده‌ی دیگر با نسبت‌های معین است. نسبت متناظر با هریک از متابولیت‌ها در هریک از واکنش‌ها را ضریب استوکیومتری آن متابولیت در واکنش مربوطه می‌گوییم. برای مثال در شکل ۲-۱ در طی واکنش  $R$  دو واحد از متابولیت  $A$  با یک واحد از متابولیت  $B$  واکنش داده و یک واحد متابولیت  $C$  حاصل می‌شود. بنابراین ضرایب استوکیومتری متابولیت‌های  $A$ ،  $B$  و  $C$  در این واکنش به ترتیب عبارتند از  $-۲$ ،  $-۱$  و  $۱$ .



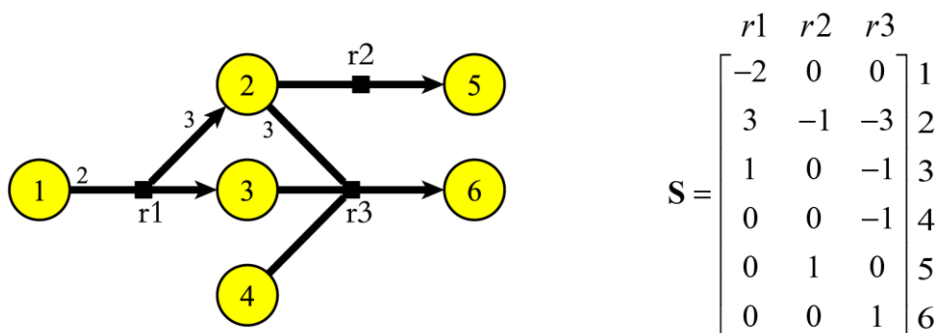
شکل ۲-۱ یک واکنش و نحوه‌ی نمایش آن

شار یک واکنش به معنی سرعت مصرف واکنش‌دهنده‌ها و تولید فراورده‌ها است. اگر شار واکنش  $R$  برابر ۲ باشد، به این معنی است که در واحد زمان ۴ واحد از متابولیت  $A$  و ۲ واحد از متابولیت  $B$  مصرف شده و ۲ واحد متابولیت  $C$  حاصل می‌شود. به طور کلی اگر ضریب استوکیومتری متابولیت  $i$  در واکنش  $j$  برابر  $S_{i,j}$  بوده و شار واکنش  $j$  برابر  $v$  باشد، سرعت تولید متابولیت  $i$  توسط واکنش  $j$  به وسیله‌ی رابطه‌ی ۲-۱ به دست می‌آید.

$$\frac{dc_i}{dt} = v \cdot S_{i,j} \quad ۲-۱$$

## ۲-۲-۲ ماتریس استوکیومتری

در یک شبکه‌ی متابولیکی متابولیت‌ها در واکنش‌های متعددی شرکت دارند و در هریک از واکنش‌ها نیز ضریب استوکیومتری به‌خصوصی دارند. معمولاً همه‌ی ضرایب استوکیومتری را در ماتریسی به نام ماتریس استوکیومتری قرار می‌دهند. این ماتریس علاوه بر نشان دادن ارتباطات متابولیت‌ها و واکنش‌ها در شبکه، ضریب استوکیومتری هر متابولیت در هر واکنش را نیز مشخص می‌سازند. سطرهای این ماتریس متناظر با متابولیت‌ها و ستون‌های این ماتریس متناظر با واکنش‌ها هستند. عنصر سطر  $i$  ام و ستون  $j$  ام، ضریب استوکیومتری متابولیت  $i$  ام را در واکنش  $j$  ام نشان می‌دهد. برای مثال در شکل ۲-۲ یک شبکه‌ی نمونه و ماتریس استوکیومتری آن نشان داده شده است.



شکل ۲-۲ یک شبکه‌ی متابولیکی و ماتریس استوکیومتری آن

## ۳-۲-۲ قیدهای حالت پایا

متابولیت‌ها اغلب همزمان در واکنش‌های متعددی شرکت می‌کنند. بنابراین یک ماده اغلب توسط یک یا چند واکنش تولید و توسط یک یا چند واکنش دیگر مصرف می‌شود. در حالت پایا فرض بر این است که غلظت مواد ثابت است. بنابراین برای هر متابولیت باید سرعت تولید برابر با سرعت مصرف باشد. این مساله موجب به وجود آمدن یک معادله‌ی جبری بین واکنش‌های متصل به یک متابولیت می‌گردد. این محدودیت‌ها را می‌توان به این صورت بیان کرد که برای هر متابولیت مجموع جبری نرخ تولید آن به ازای واکنش‌های مختلف برابر صفر است. با توجه به رابطه‌ی ۱-۲ می‌توان برای هر متابولیت  $i$  معادله‌ی ۲-۲ را نوشت.

$$\sum_{j \in R} S_{i,j} \cdot v_j = 0 \quad 2-2$$

که در آن  $S_{i,j}$  ضریب استوکیومتری واکنش  $j$ ام برای متابولیت  $i$ ام و  $v_j$  شار واکنش  $j$ ام را نشان می‌دهد. با استفاده از معادله‌ی ماتریسی ۳-۲ می‌توان معادلات حالت پایا برای همه‌ی متابولیت‌ها را به صورت یک‌جا بیان کرد.

$$S \cdot v = 0 \quad 3-2$$

که در آن  $S$  ماتریس استوکیومتری و  $v$  بردار ستونی است که هر عنصر آن شار یکی از واکنش‌ها را نشان می‌دهد.

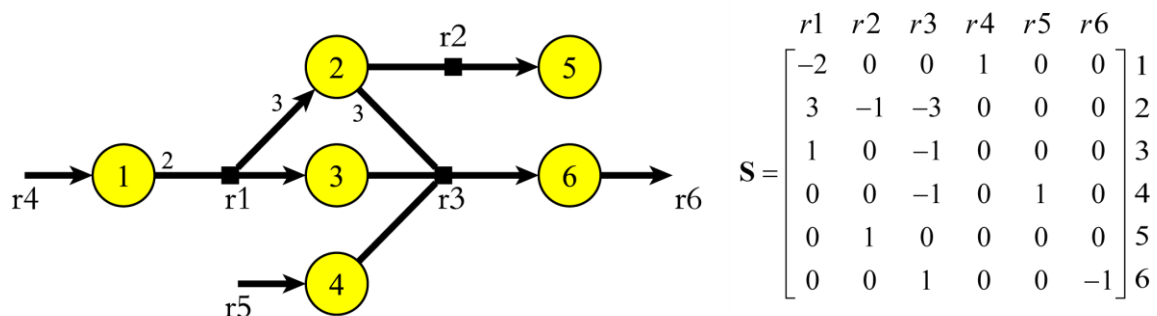
## ۲-۲-۴ واکنش‌های تبادلی و متابولیت‌های خارجی

برای این که ورودی‌ها و خروجی‌های سیستم مدل شود واکنش‌های خاصی به نام واکنش‌های تبادلی<sup>۱۹</sup> در مرزهای سیستم نظر گرفته می‌شود. وقوع این واکنش‌ها نشان دهنده‌ی یک تغییر شیمیایی نیست و ورود مواد از بیرون سیستم به داخل آن یا خروج مواد از سیستم به خارج را نشان می‌دهد. واکنش‌های تبادلی ورودی، تنها دارای فراورده و واکنش‌های تبادلی خروجی، تنها دارای واکنش دهنده هستند. بدون در نظر گرفتن چنین واکنش‌هایی در اکثر موارد هیچ شاری از درون شبکه عبور نخواهد کرد. برای مثال در شبکه‌ی شکل ۲-۲

۲-۲ هیچ یک از واکنش‌ها نمی‌توانند شار غیر صفر داشته باشند. زیرا متابولیت ۱ تنها توسط واکنش  $r1$  مصرف می‌شود و هیچ واکنشی آن را تولید نمی‌کند. واکنش‌های  $r2$  و  $r3$  نیز به دلیل مشابهی نمی‌توانند شار غیر صفر داشته باشند. با اضافه کردن واکنش‌های تبادلی  $r4$ ،  $r5$  و  $r6$  شبکه به صورت شکل ۳-۲ در می‌آید. در این شبکه‌ی جدید با اضافه شدن واکنش‌های تبادلی، شار  $v$  برای شبکه امکان‌پذیر خواهد شد.

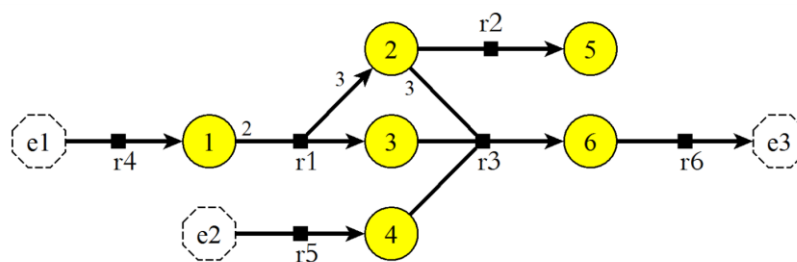
$$v = [1 \ 0 \ 1 \ 1 \ 1 \ 1]^T$$

<sup>19</sup> Exchange Reaction



شکل ۲-۳ شبکه‌ی متابولیکی با اضافه کردن واکنش‌های تبدالی

همچنین گاهی در سمت خارجی واکنش‌های تبدالی، متابولیت‌های خارجی در نظر گرفته می‌شوند. این متابولیت‌ها در معادلات حالت پایا نقشی ندارند و تنها برای این که مرزهای شبکه و ورود و خروج مواد به شبکه به شکل معنی دارتری مشخص شود، قرار داده می‌شوند. با قرار دادن این متابولیت‌های خارجی می‌توان این مفهوم را به شکل بهتری مشاهده کرد که از طریق واکنش‌های این شبکه چه موادی در نهایت به یکدیگر تبدیل می‌گردند. با اضافه کردن متابولیت‌های خارجی، شکل ۲-۴ به دست می‌آید. ضمن این که معمولاً سطرهای مربوط به متابولیت‌های خارجی را در ماتریس دیگری (ماتریس  $E$ ) نشان می‌دهند. با توجه به این که متابولیت‌های خارجی نقشی در معادلات حالت پایا ندارند، ممکن است در شبکه نشان داده نشوند.



$$E = \begin{matrix} & \begin{matrix} r1 & r2 & r3 & r4 & r5 & r6 \end{matrix} \\ \begin{bmatrix} 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} & \begin{matrix} e1 \\ e2 \\ e3 \end{matrix} \end{matrix}$$

شکل ۲-۴ شبکه‌ی متابولیکی با اضافه کردن متابولیت‌های خارجی به همراه ماتریس مربوطه

## ۵-۲-۲ واکنش‌های برگشت ناپذیر

برخی واکنش‌های شیمیایی از نظر محدودیت‌های ترمودینامیکی به گونه‌ای هستند که تنها در یک جهت می‌توانند پیشرفت کنند. به این واکنش‌ها برگشت‌ناپذیر می‌گوییم. برای این که جواب‌هایی که برای سرعت واکنش‌های یک شبکه بدست می‌آید، مقادیر قابل قبولی باشند، لازم است که شار برخی واکنش‌های داخل شبکه مثبت باشند. بنابراین قیده‌های دیگری علاوه بر قیده‌های حالت پایا به سیستم اضافه می‌شود. واکنش‌های شبکه به دو گروه برگشت پذیر ( $R_{rev}$ ) و برگشت ناپذیر ( $R_{irrev}$ ) تقسیم می‌شوند. برای شار واکنش‌های برگشت‌ناپذیر ( $\mathbf{v}_{irrev}$ ) باید نامعادله‌ی ۴-۲ برقرار باشد.

$$\mathbf{v}_{irrev} \geq \mathbf{0} \quad ۴-۲$$

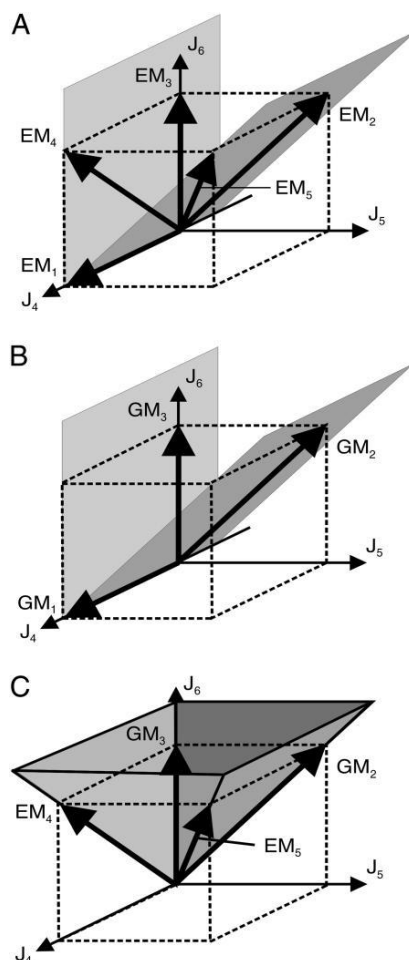
## ۶-۲-۲ فضای جواب و مخروط شار

معادلات حالت پایا به همراه نامعادلات مربوط به واکنش‌های برگشت‌ناپذیر مجموعه‌ی همه‌ی شارهای مجاز در حالات پایای شبکه‌ی متابولیکی را تعریف می‌کنند. با توجه به این که در شبکه‌های متابولیکی معمولاً تعداد واکنش‌ها (تعداد متغیرها) بسیار بیشتر از تعداد متابولیت‌ها (تعداد معادلات) است، این سیستم معمولاً نامعین بوده و دارای تعداد نامحدودی جواب می‌باشد.

مجموعه جواب بردارهای شار، یک ناحیه را در فضای  $\mathbb{R}^r$  مشخص می‌کند که به آن مخروط شار<sup>۲۰</sup> می‌گوییم. در شکل ۵-۲ می‌توان ویژگی‌های این مجموعه جواب را مشاهده کرد. این شکل نشان می‌دهد که ناحیه‌ی جواب‌های مساله، دارای یک راس در مبدا مختصات (نقطه‌ی صفر) می‌باشد و به صورت ترکیب‌های خطی مثبت (محدب<sup>۲۱</sup>) از یک مجموعه‌ی متناهی از بردارهای پایه، گسترش می‌یابند[۱].

<sup>20</sup> Flux Cone

<sup>21</sup> Convex Combintation



شکل ۲-۵ فضای جواب حاصل از معادلات و نامعادلات حاصل (مخروط شار) [۱]

## ۲-۲-۷ تحلیل‌های مربوط به مدل مبتنی بر محدودیت

مجموعه جواب حاصل از معادلات حالت پایا و شرایط برگشت‌پذیری، از جهات گوناگون می‌تواند مورد بررسی قرار بگیرد و به روش‌های مختلفی اطلاعات شهودی یا کمی از آن استخراج گردد. بر این اساس روش‌های تحلیلی متعددی برای به دست آوردن اطلاعات از این نوع مدل معرفی شده‌اند که به اختصار هر کدام را توضیح می‌دهیم.



**تحلیل شار متابولیکی<sup>۲۲</sup>:** در این روش تعدادی از شارهای قابل اندازه‌گیری به روش تجربی اندازه‌گیری می‌شوند. سپس شارهای غیرقابل اندازه‌گیری با استفاده از معادلات به دست آمده محاسبه می‌شوند. به این صورت که به اندازه‌ی کافی مقادیر معلوم جایگزین شارهای مجهول می‌شود تا معادله به صورت معین درآید و شارهای مجهول بتواند محاسبه شوند. مقادیر اندازه‌گیری شده مربوط به شرایط خاص زیستی هستند و بنابراین نتایج به دست آمده از این روش نیز برای همان شرایط معتبرند و کلیت ندارند.

**تحلیل تعادل شار<sup>۲۳</sup>:** در این روش، بیشینه یا کمینه کردن یک تابع هدف مورد توجه است. تابع هدف معمولاً یک تابع خطی از متغیرهای سیستم بوده و بنابراین بهینه‌سازی آن توسط روش‌های برنامه‌ریزی خطی قابل انجام است. این تابع معمولاً با استفاده از یک شهود زیستی یا یک انگیزه‌ی کاربردی تعریف می‌شود. برای مثال یک شهود زیستی در مورد موجودات تک‌سلولی این است که این موجودات طوری تکامل یافته‌اند که سعی در حداکثر کردن رشد و تولید مثل خود دارند و بنابراین متابولیسم خود را همواره طوری تنظیم می‌نمایند که تولید مواد ساختاری برای رشد و تولید مثل بیشینه شود. در تحقیقاتی با محاسبه‌ی حالت سیستم در این شرایط بهینه و مقایسه‌ی آن با اندازه‌گیری‌های آزمایشگاهی، این شهود مورد آزمایش قرار گرفته و نشان داده شده است که بر مشاهدات تجربی انطباق دارد [۸]. همچنین ممکن است مثلاً در یک مورد کاربردی همانند ساخت سوخت زیستی هدف بیشینه کردن تولید یک ماده‌ی خاص باشد. در این صورت با حل مساله‌ی بهینه‌سازی به وجود آمده می‌توان حالتی از سیستم که شرایط مطلوب را به وجود می‌آورد به دست آورد و سپس متابولیسم موجود را با استفاده از دستکاری‌های مختلف به سمت مطلوب سوق داد.

**تحلیل مسیرهای متابولیکی<sup>۲۴</sup>:** در این روش به جای بررسی یک حالت خاص سیستم، به بررسی همه‌ی فضای جواب از طریق معرفی یک مجموعه‌ی متناهی از بردارها پرداخته می‌شود، با این هدف که این بردارها بتوانند فضای حالات ممکن را به گویاترین و خلاصه‌ترین شکل توصیف نمایند. این بردارهای جواب، محدود به شرایط خاص بیولوژیکی نیستند و ویژگی‌های سیستم را در حالت کلی مشخص

<sup>22</sup> Metabolic Flux Analysis

<sup>23</sup> Flux Balance Analysis

<sup>24</sup> Metabolic Pathway Analysis

می‌کنند. داشتن این بردارهای پایه، همچنین امکان برخی پیش‌بینی‌ها و دستکاری‌های هدف‌مند را فراهم می‌کنند. مدهای پایه‌ی شار<sup>۲۵</sup> که به عنوان یک تعریف دقیق ریاضی برای مفهوم سنتی مسیر متابولیکی پذیرفته شده است را می‌توان پرکاربردترین ابزار در این زمینه دانست. هرچند مفاهیم مرتبط دیگری همچون مسیرهای کرانی، مسیرهای مولد، و رفتار کمینه‌ی متابولیکی<sup>۲۶</sup> نیز در این زمینه تعریف شده‌اند. مشکلی که در زمینه‌ی تحلیل مسیرهای متابولیکی وجود دارد، مشکل ذاتی این روش‌ها با شبکه‌های بزرگ است. با وجود این که در مورد شبکه‌های کوچک تعداد بردارهای پایه محدود است و به درک شهودی از قابلیت‌های یک شبکه‌ی متابولیکی کمک می‌کند، در مورد شبکه‌های بزرگ مقیاس ژنوم تعداد این بردارها بسیار بیشتر از تعداد متابولیت‌ها و واکنش‌ها است. برای مثال در مورد شبکه‌های کنونی با حدود ۱۰۰۰ متابولیت و واکنش، تعداد مسیرها ممکن است بیش از چند میلیارد باشد. این مشکل که به انفجار ترکیبیاتی<sup>۲۷</sup> معروف است، باعث شده که اعمال مستقیم روش مسیرهای متابولیکی بر روی شبکه‌های بزرگ عملی نباشد.

**تحلیل کوپلینگ شار<sup>۲۸</sup>:** قیدهای ایجاد شده بر روی شار واکنش‌ها می‌تواند موجب شود که بین بعضی از واکنش‌ها وابستگی‌های مستقیم ایجاد شود. برای مثال در شبکه‌ی شکل ۲-۴ مقدار شار دو واکنش  $r1$  و  $r3$  همواره با هم برابرند (به دلیل معادله‌ی متناظر با متابولیت ۳). همچنین شار واکنش  $r4$  همواره دو برابر شار واکنش  $r5$  است. این نوع کوپلینگ را کوپلینگ کامل می‌گوییم. روابط کوپلینگ شامل سه نوع رابطه‌ی زیر می‌باشند [۹].

(۱) کوپلینگ یک طرفه: هنگامی که وجود شار برای واکنش  $i$  وجود شار برای واکنش  $j$  را نتیجه بدهد. اما لزوماً عکس آن صادق نباشد. ( $i \rightarrow j$ )

(۲) کوپلینگ جزئی: هنگامی که وجود شار برای واکنش  $i$ ، وجود شار برای واکنش  $j$  را نتیجه می‌دهد و بالعکس. ( $i \leftrightarrow j$ )

<sup>25</sup> Elementary Flux Modes

<sup>26</sup> Minimal Metabolic Behaviour

<sup>27</sup> Combinatorial Explosion

<sup>28</sup> Flux Coupling Analysis

۳) کوپلینگ کامل: هنگامی که شار واکنش  $i$  همواره ضریب ثابتی از شار واکنش  $j$  است ( $i \Leftrightarrow j$ ) تحلیل کوپلینگ شار بستگی‌های بین هر زوج از واکنش‌ها را بررسی می‌کند، به این صورت که معمولاً ماتریسی تشکیل می‌شود که رابطه‌ی کوپلینگ هر واکنش  $i$  با هر واکنش دیگر  $j$  در آن مشخص می‌شود. این روش مقیاس‌پذیری خوبی برای اعمال بر روی شبکه‌های بزرگ دارد. در مقابل نقطه ضعف این روش این است که در مقابل نقص‌هایی که ممکن است در ساخت مدل وجود داشته باشد، حساس است. بنابراین قبل از استفاده از نتایج این تحلیل باید از کامل بودن شبکه اطمینان حاصل نمود [۱۰، ۱۱].

**تحلیل تغییرپذیری شار<sup>۲۹</sup>:** در این روش، با توجه به فضای جواب ناشی از قیدهای خطی، بازه‌ی ممکن برای تغییرات شار یک واکنش خاص مشخص می‌شود. این کار می‌تواند به راحتی با استفاده از برنامه‌ریزی خطی یک بار با هدف کمینه کردن شار واکنش مورد نظر و یک بار با هدف بیشینه کردن شار واکنش، انجام گیرد. با استفاده از این تحلیل ضمن به دست آوردن اطلاعات مفید در مورد تک واکنش‌ها، می‌توان واکنش‌هایی که همواره شار ثابتی دارند را مشخص کرد. چنین واکنش‌هایی دارای بیشینه و کمینه‌ی شار یکسانی خواهند بود [۱۲].

## ۲-۳ مدهای پایه‌ی شار

یکی از مهم‌ترین تحلیل‌های انجام شده در مدل مبتنی بر محدودیت در شبکه‌های متابولیکی تحلیل با استفاده از مدهای پایه‌ی شار است. این تحلیل که در دسته‌ی تحلیل مسیرهای متابولیکی قرار دارد، به معرفی یک مجموعه‌ی متناهی و یکتا از بردارهای شار می‌پردازد که دارای حداقل واکنش‌های ممکن برای فعالیت در شرایط حالت پایا هستند. برای روشن تر شدن این مفهوم تعریف دقیق مدهای پایه‌ی شار ارائه می‌گردد.

<sup>29</sup> Flux Variability Analysis

**تعریف (مد پایه‌ی شار):** فرض می‌کنیم شبکه‌ی متابولیکی  $N$  دارای ماتریس استوکیومتری  $S$  باشد و  $irrev$ ، اندیس واکنش‌های برگشت ناپذیر را نشان دهد، مجموعه‌ی جواب محدودیت‌های حالت پایا و برگشت‌ناپذیری واکنش‌ها یا به عبارتی مخروط شار به صورت زیر خواهد بود.

$$C = \{\mathbf{x} \in \mathbb{R}^r : \mathbf{S} \cdot \mathbf{x} = \mathbf{0}, \mathbf{x}_{irrev} \geq \mathbf{0}\} \quad ۵-۲$$

در این صورت بردار  $\mathbf{e}$  یک مد پایه‌ی شار برای شبکه‌ی  $N$  است اگر و فقط اگر هر دو شرط زیر برقرار باشد:

$$\mathbf{e} \in C \quad (۱)$$

$$\forall \mathbf{x} \in C : \text{supp}(\mathbf{x}) \subseteq \text{supp}(\mathbf{e}) \Rightarrow \mathbf{x} = \lambda \mathbf{e} \quad \lambda \in \mathbb{R} \quad (۲)$$

که  $\text{supp}$  مجموعه‌ی اندیس‌هایی از یک بردار را نشان می‌دهد که مقدار غیر صفر دارند.

قسمت دوم تعریف مدهای پایه این را بیان می‌کند که اگر مجموعه‌ی واکنش‌های (غیر صفر) یک بردار شار مجاز، زیر مجموعه‌ای از واکنش‌های (غیر صفر) یک مد پایه باشد، در آن صورت این بردار حتماً ضربی از آن مد پایه خواهد بود. به عبارت دیگر مدهای پایه مدهایی هستند که مجموعه‌ی واکنش‌های غیر صفر آن‌ها غیر قابل کاهش است و هیچ بردار شار مجازی وجود ندارد که مجموعه‌ی واکنش‌های آن زیرمجموعه‌ی سرهای از مجموعه‌ی واکنش‌های آن‌ها باشد.

تعریف شهودی مدهای پایه به صورت مسیریایی از واکنش‌ها است که حداقل واکنش‌های ممکن را دارند و با حذف یک یا چند واکنش از آن‌ها، حتماً یک مسیر غیر مجاز (از لحاظ شرط حالت پایا) حاصل خواهد شد.

## ۴-۲ روش توصیف دوتایی برای محاسبه‌ی مدهای پایه شار

ثابت شده است که اگر همه‌ی واکنش‌های شبکه برگشت‌ناپذیر باشند، در آن صورت مدهای پایه برابر شعاع‌های کرانی<sup>۳۰</sup> (لبه‌های) ناحیه‌ی چند وجهی محدبی<sup>۳۱</sup> هستند که توسط معادلات استوکیومتری و

<sup>30</sup> Extreme Rays

نامعادلات برگشت ناپذیری تعریف می‌شوند [۱۳]. در هندسه‌ی محاسباتی<sup>۳۲</sup> برای به دست آوردن این لبه‌ها از روشی به نام توصیف دوتایی<sup>۳۳</sup> استفاده می‌شود. این روش شامل مراحل زیر است.

(۱) ابتدا از یک ناحیه شروع می‌شود که تعدادی از قیدهای تعریف‌کننده‌ی ناحیه را برآورده می‌سازند. برای این ناحیه‌ی اولیه، لبه‌های ناحیه به دست می‌آیند.

(۲) در هر مرحله یکی از شرط‌های باقی‌مانده به ناحیه‌ی فعلی اعمال می‌شود و بخشی از ناحیه حذف گردد. بنابراین تعدادی از لبه‌ها حذف و تعدادی لبه‌های جدید به وجود می‌آیند. در هر مرحله برای به وجود آمدن لبه‌های جدید، از ترکیب لبه‌های قبلی چندین کاندیدا به وجود می‌آید. تنها کاندیداهایی به عنوان لبه پذیرفته می‌شوند، که از ترکیب لبه‌های مجاور به وجود آمده باشند. بنابراین برای به دست آوردن لبه‌ها در هر مرحله ابتدا باید لبه‌های مجاور را مشخص کرد و سپس از ترکیب آن‌ها لبه‌های جدید را به دست آورد. سپس از میان لبه‌های مرحله‌ی قبل آن‌هایی که شرط جدید را برآورده نمی‌سازند حذف می‌گردند. حاصل این مرحله لبه‌های یک ناحیه است که همه‌ی شرط‌های مراحل قبلی به علاوه‌ی یک شرط جدید را برآورده می‌سازد.

(۳) مرحله‌ی ۲ آن قدر تکرار می‌شود تا همه‌ی شرط‌ها اعمال گردند. در نتیجه لبه‌های ناحیه‌ای به دست می‌آید که همه‌ی شرط‌های داده شده را برآورده می‌سازد.

در الگوریتم پوچ‌فضا<sup>۳۴</sup> که یکی از الگوریتم‌های مبنی بر روش توصیف دوتایی است، ابتدا از ناحیه‌ای شروع می‌کنیم که همه‌ی تساوی‌ها را برآورده می‌کند. بردارهای پایه‌ی این ناحیه به راحتی از روی ماتریس کرنل<sup>۳۵</sup> به دست می‌آیند. سپس نامساوی‌ها یک به یک به ناحیه اعمال می‌شوند.

با توجه به این که در شبکه‌های متابولیکی بعضی از واکنش‌ها (و نه همه‌ی آن‌ها) برگشت‌ناپذیرند بنابراین واکنش‌های برگشت‌پذیر باید به دو واکنش برگشت‌ناپذیر (در جهت‌های عکس یکدیگر) تبدیل شوند. به

<sup>31</sup> Convex Polyhedra

<sup>32</sup> Computational Geometry

<sup>33</sup> Double Description Method

<sup>34</sup> Null Space

<sup>35</sup> Kernel (Null Space) Matrix

این ترتیب یک شبکه‌ی تغییر یافته به دست می‌آید. با به دست آوردن لبه‌های ناحیه‌ی جواب برای شبکه‌ی تغییر یافته و یک مرحله‌ی پس‌پردازش، می‌توان مدهای پایه را برای شبکه‌ی اصلی محاسبه نمود [۱۳].

نکته‌ی اصلی در مورد روش‌های مبتنی بر روش توصیف دوتایی این است که این روش‌ها شامل یک حلقه‌ی اصلی هستند که در هر مرحله از ترکیب نتایج مرحله‌ی قبل تعدادی کاندیدا تولید می‌شوند که باید تعدادی از آن‌ها حذف و تعدادی تایید شوند. عملیات بررسی بردارها معمولاً توسط یکی از روش‌های زیر انجام می‌گیرد:

(۱) روش آزمون مجاورت<sup>۳۶</sup>: در این روش هر جفت از بردارهای به دست آمده از مرحله‌ی قبل با یکدیگر مقایسه شوند. این مقایسه ساده است و با یک عملیات منطقی قابل انجام است. اما به دلیل این که هر دو برداری باید با یکدیگر مقایسه گردند ممکن است تعداد مقایسه‌ها خیلی زیاد باشد. این روش معمولاً برای پیاده‌سازی‌های انجام شده به صورت حافظه‌ی اشتراکی<sup>۳۷</sup> مانند ابزار efmtool مناسب است.

(۲) روش آزمون رتبه<sup>۳۸</sup>: در این روش با ترکیب بردارهای به دست آمده، تعدادی کاندیدا تولید می‌شود که می‌توانند مستقل از یکدیگر مورد آزمون قرار بگیرند. با استفاده از محاسبه‌ی رتبه‌ی یک زیرماتریس از ماتریس استوکیومتری (شامل ستون‌هایی از ماتریس استوکیومتری که واکنش متناظر آن‌ها در بردار تحت آزمون غیر صفر است) می‌توان مشخص نمود که آیا بردار مورد نظر می‌تواند یک لبه باشد یا خیر. این روش شامل محاسبه‌ی رتبه‌ی ماتریس است. بنابراین از نظر محاسباتی سنگین‌تر از مقایسه‌ی موجود در روش ۱ است. ضمن این که به دلیل ماهیت عددی روش برای خطاهای عددی مستعدتر می‌باشد. اما به دلیل این که در این روش هریک از کاندیداها می‌توانند مستقلاً مورد آزمون قرار بگیرند، این روش برای پیاده‌سازی‌های موازی روی سیستم‌های با حافظه‌ی توزیع شده<sup>۳۹</sup> مناسب‌تر است [۱۴]. این روش در پیاده‌سازی metatool به کار گرفته شده است [۱۵].

<sup>36</sup> Adjacency Test

<sup>37</sup> Shared Memory

<sup>38</sup> Rank Test

<sup>39</sup> Distributed Memory

۳

فصل سوم

پیشینه پژوهش

## پیشینه پژوهش

### ۳-۱ مدهای پایه‌ی شار و محاسبه

تعریف مدهای پایه‌ی شار اولین بار توسط شوستر<sup>۱</sup> ارائه شده است [۱۶]. این بردارها ضمن این که به خوبی مفهوم مسیر متابولیکی را به صورت ریاضی بیان می‌کنند [۱۷]، برای کشف مسیرهای جدید متابولیکی و کاربردهای فراوان دیگر به کار رفته‌اند [۱۸، ۱۹]. با توجه به کاربرد و اهمیت فراوان آن‌ها تلاش‌های زیادی نیز به منظور محاسبه‌ی بهینه‌ی این بردارها انجام گرفته است [۱۹، ۲۰].

### ۳-۱-۱ روش‌های مبتنی بر هندسه‌ی محاسباتی

یکی از اولین الگوریتم‌های محاسبه‌ی مدهای پایه الگوریتم پایه‌ی کانونی<sup>۲</sup> بوده که توسط شوستر ارائه شده است. الگوریتم دیگر یعنی الگوریتم پوچ‌فضا، که بعد از آن توسط واگنر<sup>۳</sup> [۲۱، ۲۲] معرفی شده، کارایی بهتری از روش قبلی نشان داده است. پس از ارائه‌ی این الگوریتم‌ها، این نکته توسط کلمت<sup>۴</sup> بیان شده است که هر دوی آن‌ها در واقع مبتنی بر روش توصیف دوتایی در هندسه‌ی محاسباتی بوده و شعاع‌های کرانی یک چند وجهی را در فضا برمی‌شمارند [۱۳]. همچنین اکثر الگوریتم‌ها و ابزارهای ارایه شده‌ی بعدی بر مبنای الگوریتم پوچ‌فضا و با هدف ایجاد بهبود در مراحل مختلف این الگوریتم به وجود آمده‌اند.

### ۳-۱-۲ ابزارهای اولیه و بهبودهای ایجاد شده

ابزار Metatool از اولین ابزارهایی بوده که برای محاسبه‌ی مدهای پایه به کار رفته است [۲۳]. این برنامه در ابتدا پیاده‌سازی الگوریتم پایه‌ی کانونی بوده که توسط شوستر ارائه شده و بعداً بهبودهایی بر روی آن

<sup>۱</sup> Schuster

<sup>۲</sup> Canonical Basis

<sup>۳</sup> Wagner

<sup>۴</sup> Klamt



صورت گرفته است. پس از آن بهبودهای متعددی برای الگوریتم پوچ فضا پیشنهاد شده و در ابزار Metatool پیاده‌سازی گردیده است [۱۴]. برای مثال این نکته مشخص شده که یکی از ویژگی‌های شاخص مدهای پایه این است که با مشخص شدن واکنش‌های غیر صفر آن‌ها، می‌توان مقادیر شار آن‌ها را به طور یکتایی محاسبه نمود. یعنی برای مشخص کردن یک مد پایه لازم نیست که همه‌ی مقادیر شار داده شود. بلکه می‌توان آن‌ها را به صورت بردارهای دودویی ذخیره یا بیان کرد. به این صورت که اگر واکنش مربوطه دارای شار صفر باشد، در بردار دودویی با مقدار صفر و اگر دارای شار غیر صفر باشد، با مقدار یک نشان داده می‌شود. این یافته باعث بهبود چشمگیری در محاسبه‌ی مدهای پایه می‌گردد زیرا نه تنها می‌توان مدهای پایه را در مراحل میانی الگوریتم به صورت دودویی در حافظه ذخیره کرد بلکه می‌توان برخی عملیات بر روی آن‌ها را نیز به صورت دودویی انجام داد [۱۳].

همچنین در ادامه بهبودهای دیگری در الگوریتم پوچ فضا معرفی شده است. یکی از این روش‌ها استفاده از آزمون رتبه‌ی<sup>۱</sup> ماتریس برای بررسی بردارهای میانی به دست آمده است. الگوریتم پوچ فضا شامل حلقه‌ای است که در هر پیمایش حلقه، کاندیداهای جدیدی تولید می‌شوند. تعدادی از این کاندیداها پذیرفته و تعدادی از آن‌ها حذف می‌شوند. در الگوریتم اولیه‌ی پوچ فضا یک کاندیدای به دست آمده به منظور آزمون مجاورت، باید با همه‌ی بردارهای به دست آمده از مرحله‌ی قبل مقایسه گردد. اما کلمت یک روش بر مبنای آزمون رتبه‌ی ماتریس ارایه کرد که در آن به جای مقایسه‌ی هر با بردارهای به دست آمده‌ی قبلی (که ممکن است تعداد آن‌ها خیلی زیاد باشد) به تنهایی می‌تواند مورد آزمون قرار گیرد. در این روش به ازای هر بردار کاندیدا باید رتبه‌ی یک زیرماتریس از ماتریس استوکيومتری محاسبه شود [۱۴].

هرچند عملیات محاسبه‌ی رتبه‌ی یک ماتریس می‌تواند وقت‌گیرتر از مقایسه‌ی دو بردار بایکدیگر باشد، اما با توجه به این که عملیات تست رتبه برای هر کاندیدا می‌تواند مستقلاً انجام گیرد، زمان لازم برای تست کاندیدا مستقل از تعداد بردارهای به دست آمده از مرحله‌ی قبل است. همچنین به دلیل استقلال عملیات، تست این کاندیداها می‌تواند به راحتی به صورت موازی انجام گیرد. بر اساس همین رویکرد پیاده‌سازی موازی این الگوریتم با استفاده از روش حافظه‌ی توزیع‌شده با استفاده از کتابخانه‌ی تبادل

---

<sup>1</sup> Rank Test

پیام<sup>۱</sup> انجام گرفته و با اجرای آن بر روی دو نوع بستر محاسبات با کارایی بالا<sup>۲</sup> چندین میلیون مدپایه برای یک شبکه‌ی نسبتاً بزرگ به دست آمده است [۲۴]. در یک رویکرد متفاوت توسط ترزر<sup>۳</sup> به جای استفاده از تست رتبه‌ی ماتریس، تست مجاورت بردارها با استفاده از یک ساختمان داده‌ی بهینه شده به نام درخت‌های الگوی بیتی<sup>۴</sup> انجام می‌گیرد. این روش باعث پیشرفت قابل توجهی در سرعت محاسبه‌ی مدهای پایه گشته و امکان محاسبه‌ی مدهای پایه‌ی شبکه‌های بزرگتری را برای کامپیوترهای معمولی فراهم آورده است. این روش در ابزار efmtool به کار رفته که هم اکنون سریع‌ترین ابزار برای محاسبه‌ی مدهای پایه بر روی یک کامپیوتر منفرد به شمار می‌آید [۲۵].

در مقاله‌ی کلمت همچنین یک روش تقسیم و غلبه<sup>۵</sup> برای محاسبه‌ی مدهای پایه ارائه شده که در آن مساله‌ی محاسبه‌ی مدهای پایه بر اساس وجود یا عدم وجود یک واکنش خاص به دو زیر مساله تقسیم می‌شود [۱۴]. هریک از این زیرمساله‌ها می‌تواند مستقلاً حل شده و از اجتماع جواب‌های آن‌ها جواب مساله‌ی اصلی به دست آید. بر مبنای این روش نیز پیاده‌سازی‌های موازی با استفاده از بسترهای محاسبات با کارایی بالا انجام گرفته و چندین میلیون مد پایه برای چند شبکه‌ی بزرگ محاسبه شده‌اند [۲۶، ۲۷].

### ۳-۱-۳ استفاده از روش‌های مبتنی بر گراف

با توجه به این که مدهای پایه مسیرهای حداقل را در شبکه نشان می‌دهند، در تحقیقات متعددی سعی شده که مفهوم مدهای پایه با استفاده از نظریه‌ی گراف تعریف و محاسبه گردد. برای نمونه در یک روش مبتنی بر گراف، ابتدا با استفاده از الگوریتم دیکسترا<sup>۶</sup> کوتاهترین مسیر بین دو گره یافته می‌شود سپس یال‌های لازم دیگر به مسیر اضافه می‌شوند [۲۸].

<sup>۱</sup> Message Passing Interface (MPI)

<sup>۲</sup> High Performance Computing (HPC) Platforms

<sup>۳</sup> Terzer

<sup>۴</sup> Bit Pattern Tree

<sup>۵</sup> Divide and Conquer

<sup>۶</sup> Dijkstra

همچنین در یک روش دیگر در گراف شبکه با استفاده از اطلاعات مربوط به انرژی آزاد گیبس<sup>۱</sup> به هر یک از یال‌های شبکه وزن داده شده است. و هدف پیدا کردن مسیری است که از لحاظ انرژی ارجحیت دارد. وزن یک مسیر را گلوگاه انرژی در طول مسیر تعیین می‌کند و این گلوگاه واکنشی است که کمترین تغییرات انرژی آزاد گیبس را دارد. به این طریق، هدف این الگوریتم پیدا کردن مسیرهایی است که برای تولید یک متابولیت خاص، از لحاظ انرژی، گلوگاه کمتری داشته باشند و به این ترتیب بر مسیرهای دیگر ترجیح داشته باشند. به منظور کامل کردن مسیر به دست آمده برای ایجاد تعادل استوکیومتری گره‌هایی که به عنوان شاخه‌های جانبی واکنش‌های موجود در مسیر ایجاد شده‌اند، به متابولیت‌های خارجی متصل می‌شوند [۲۹].

اخیراً در تحقیقی با استفاده از پیمایش گراف، الگوریتمی برای یافتن همه‌ی مدهای پایه‌ی بین دوتا از گره‌ها ارائه شده که کارایی آن در مواردی بهتر و در مواردی بدتر از الگوریتم‌های موجود بوده است [۷].

### ۳-۱-۴ پیچیدگی زمانی مسائل مربوط به مدهای پایه

با وجود تحقیقات بسیار بر روی محاسبه‌ی مدهای پایه و الگوریتم‌های محاسبه‌ی آن‌ها پیچیدگی زمانی مساله‌ی محاسبه‌ی مدهای پایه هنوز ناشناخته است. هر چند پیچیدگی چندین مساله در ارتباط با مدهای پایه به دست آمده است. برای مثال:

۱. پیدا کردن یک مد پایه برای یک شبکه‌ی متابولیکی می‌تواند در زمان چند جمله‌ای انجام گیرد.
۲. مساله‌ی محاسبه‌ی همه‌ی مدهای پایه‌ای که شامل یک مجموعه واکنش  $F$  و فاقد یک مجموعه واکنش  $B$  باشند، می‌تواند در زمان چند جمله‌ای نسبت به اندازه‌ی خروجی انجام گیرد اگر و فقط اگر  $|B|=1$ . اندازه‌ی مجموعه‌ی  $B$  بر پیچیدگی این مساله نمی‌افزاید زیرا این مساله را می‌توان با حذف واکنش‌های مجموعه‌ی  $B$  از شبکه به دست آورد [۳۰، ۳۱].

<sup>1</sup> Gibbs Free Energy

### ۳-۱-۵ محاسبه‌ی زیرمجموعه‌ای از مدهای پایه

ابزارهای اصلی محاسبه‌ی مدهای پایه قابلیت محاسبه‌ی تدریجی مجموعه‌ی مدهای پایه را ندارند. یعنی نمی‌توان بخشی از مدهای پایه را قبل از پایان اجرای الگوریتم به دست آورد. بلکه برای به دست آوردن هر تعداد مد پایه باید تا پایان اجرای این الگوریتم‌ها صبر کرد. با توجه به اندازه‌ی بزرگ شبکه‌های متابولیکی مقیاس ژنوم و تعداد بسیار زیاد مدهای پایه در این شبکه‌ها امکان محاسبه‌ی همه‌ی آن‌ها وجود ندارد. بنابراین برخی از تحقیقات به ارائه‌ی روش‌هایی برای محاسبه‌ی بخشی از مدهای پایه‌ی شبکه پرداخته‌اند.

در یک پژوهش، روشی با ماهیت متفاوتی از الگوریتم‌های موجود، برای یافتن کوتاهترین مدهای پایه ارائه شده است. به این صورت که با استفاده از برنامه‌ریزی خطی صحیح<sup>۱</sup>، مدهای پایه‌ای که دارای کمترین تعداد واکنش هستند به ترتیب صعودی تعداد واکنش‌ها محاسبه می‌شوند. مزیت این روش این است که با توجه به این که محاسبه‌ی مدهای پایه تدریجی است، می‌توان در هر مرحله عملیات را قطع و تعدادی مد پایه را به دست آورد. البته با توجه به این که این روش از برنامه‌ریزی خطی صحیح استفاده می‌کند، با زیاد شدن تعداد واکنش‌ها، این روش به شدت کند می‌شود. در نتیجه از حدی به بعد نمی‌توان مدهای پایه‌ی بیشتری محاسبه کرد. اهمیت این روش به دلیل قابلیت محاسبه‌ی تعدادی هرچند محدود از مدهای پایه برای شبکه‌های بزرگ است که قبل از آن امکان پذیر نبوده است. همچنین این نکته بیان شده است که کوتاهترین مدهای پایه دارای اهمیت‌های زیست‌شناختی نیز هستند [۳۲].

در مقابل، در تحقیقاتی دیگر [۳۳] بر لزوم تصادفی بودن نمونه برداری از مدهای پایه‌ی شبکه تاکید شده است و بر این اساس، یک روش نمونه برداری بر اساس گنجانیدن یک مرحله‌ی تصادفی در الگوریتم‌های موجود معرفی شده است. در الگوریتم‌های موجود که بر اساس روش توصیف دوتایی عمل می‌کنند، یک حلقه‌ی اصلی وجود دارد که در هر پیمایش حلقه، از ترکیب بردارهای به دست آمده از مرحله‌ی قبل، تعدادی کاندیدا تولید می‌شوند. هنگامی که تعداد بردارهای به دست آمده از مرحله‌ی قبل زیاد باشد، تعداد حالت‌های ترکیب این بردارها با یکدیگر بسیار زیاد خواهد شد. که در نتیجه‌ی آن مشکل انفجار ترکیبیاتی به وجود می‌آید. برای جلوگیری از این مشکل، یک فیلتر تصادفی بر روی فرایند تولید

<sup>۱</sup> Integer Linear Programming

کاندیدها قرار داده شده تا در هر مرحله به جای تولید همه‌ی کاندیدها، به طور تصادفی تعدادی از آن‌ها تولید شوند. این انتخاب تصادفی بر مبنای یک توزیع احتمالی با یک پارامتر قابل تعیین رخ می‌دهد که این پارامتر تعیین می‌کند که چه تعداد بردار کاندیدا تولید شده و چه تعداد نادیده گرفته شوند. بنابراین این پارامتر باید طوری تنظیم شود که از طرفی مانع از به وجود آمدن مشکل انفجار ترکیباتی شود و از طرف دیگر موجب از دست رفتن همه‌ی مدهای پایه نگردد. در واقع هدف به دست آوردن تعداد معقولی از مدهای پایه‌ی شبکه است که به دلیل انتخاب تصادفی آن‌ها می‌توانند بیانگر ویژگی‌های مجموعه‌ی کلی مدهای پایه باشند.

در ادامه در تحقیقات متعددی سعی شده است که اطلاعات و اهداف زیستی وارد رویکرد مبتنی بر نمونه‌برداری شود. بر همین مبنا در پژوهش دیگری با استفاده از الگوریتم ژنتیک به یافتن مدهای پایه در شبکه‌های متابولیکی پرداخته می‌شود. به این صورت که واکنش‌هایی به طور تصادفی از شبکه حذف می‌شوند. پس از آن، یک مد پایه برای شبکه‌ی کاهش یافته محاسبه می‌شود. با استفاده از الگوریتم ژنتیک این روند برای بهینه کردن یک هدف زیستی هدایت می‌شود [۳۴]. در ادامه‌ی همین کار در تحقیق دیگری برای بهبود عملکرد این روش از یک روش درختی به همراه برنامه‌ریزی خطی استفاده شده است [۳۵].

یکی دیگر از رویکردها به دست آوردن مدهای پایه‌ای که دارای تعدادی شرایط اضافی تعیین شده هستند می‌باشد. برای مثال روشی برای یافتن مدهای پایه‌ای که دارای تعدادی قیدهای اضافه به صورت نامعادله‌ی خطی هستند ارائه شده است [۳۶]. همچنین روشی برای یافتن تعداد مشخصی از مدهای پایه که شامل یک مجموعه واکنش داده شده هستند مطرح شده است [۳۷].

یکی از رویکردهای جدید و مهم در زمینه‌ی مدهای پایه، استفاده از داده‌های زیستی دیگر به منظور عدم محاسبه‌ی مسیرهایی است که از لحاظ زیستی یا ترمودینامیکی غیر قابل قبول و بی‌معنی هستند. مدهای پایه‌ای که از یک شبکه‌ی متابولیکی به دست می‌آیند، تنها با استفاده از اطلاعات ماتریس استوکیومتری و تعیین واکنش‌های برگشت‌ناپذیر به دست آمده‌اند. ممکن است مسیرهایی به دست بیایند که از لحاظ توپولوژیکی و استوکیومتری صحیح اما از لحاظ زیست شناختی بی‌معنی باشند. برای مثال برای این که یک مسیر متابولیکی بتواند به صورت خودبه‌خودی جریان داشته باشد، باید انرژی آزاد

گیس فرایند متناظر با آن منفی باشد. در یک تحقیق با استفاده از داده‌های زیستی متابولومیک<sup>۱</sup> این واقعیت به دست آمده است که ۸۰ درصد مسیرهایی که برای یک شبکه‌ی متابولیکی با استفاده از روش مدهای پایه و مستقیماً از شبکه‌ی استوکیومتری به دست آمده‌اند، از لحاظ ترمودینامیکی و انرژی آزاد گیبس غیر قابل قبول بوده‌اند [۳۸]. در راستای همین مشاهدات، روشی ارایه شده است که در آن در فرایند محاسبه‌ی مدهای پایه از اطلاعات ترمودینامیکی موجود در داده‌های زیستی استفاده می‌شود و به این صورت بدون از دست رفتن مسیرهای معنی دار، از محاسبه‌ی مسیرهایی که از لحاظ ترمودینامیکی بی‌معنی هستند جلوگیری می‌شود. با این کار همچنین از بار عملیاتی روش کاسته می‌شود و امکان محاسبه‌ی مدهای پایه برای شبکه‌های بزرگتری فراهم می‌شود [۳۹].

همچنین یک واقعیت زیستی دیگر حاکی از این است که بسیاری از واکنش‌های موجود در شبکه‌های متابولیکی به دلیل کنترل‌های ژنتیکی نمی‌توانند همزمان رخ دهند. دلیل این مطلب این است که ژن‌هایی که آنزیم مربوط به این واکنش‌ها را می‌سازند در هیچ شرایطی به طور همزمان فعال نخواهند بود. این وابستگی‌های منطقی بین ژن‌ها که وابستگی‌های منطقی بین واکنش‌ها را نیز در پی دارند، در شبکه‌هایی به نام شبکه‌های تنظیم ژن<sup>۲</sup> فرموله شده و در پایگاه داده‌های زیستی موجودند. گروهی از این واقعیت استفاده کرده‌اند تا با استفاده از اطلاعات موجود و گنجاندن این وابستگی‌های در فرایند به دست آوردن مدهای پایه از محاسبه‌ی مدهای پایه‌ای که از لحاظ زیستی هرگز رخ نخواهند داد جلوگیری کنند. این روش در نرم افزار efmtool گنجانده و بر روی شبکه‌های زیستی امتحان شده است. این روش نیز بدون از دست رفتن مسیرهای با معنا باعث حذف بخشی از جواب‌های اضافه و کاهش زمان اجرا و حافظه‌ی مصرفی نسبت به روش‌های عادی شده‌اند [۴۰].

استفاده از داده‌های زیستی، رویکردی برای حذف بردارهای بی‌معنی از مدهای پایه‌ی محاسبه شده هستند. هرچند فرایند پردازش و فیلتر کردن مدهای پایه با استفاده از داده‌های زیستی، می‌تواند بعد از محاسبه‌ی مجموعه‌ی کامل مسیرهای توپولوژیکی نیز انجام گیرد. اما حذف زود هنگام این بردارها

<sup>۱</sup> Metabolomic

<sup>۲</sup> Gene Regulation Network

می‌تواند با کاهش بار عملیاتی در فرایند محاسبه مشکل حافظه و زمان اجرا را نیز تا حدودی بهبود دهد. و امکان محاسبه‌ی مدهای پایه‌ی معنی دار را برای شبکه‌های بزرگتری فراهم آورد.

در رویکرد متفاوتی، گروهی با تاکید بر ویژگی مولد بودن مدهای پایه سعی کرده‌اند مجموعه‌ی متفاوتی از بردارهای مولد را به دست آورند که از نظر تعداد کمتر از مدهای پایه هستند. مدهای پایه بردارهایی هستند که ضمن این که همه‌ی مسیرهای استوکیومتری متعادل و غیرقابل تجزیه در شبکه‌های متابولیکی را نشان می‌دهند، دارای این ویژگی مهم هستند که از ترکیب خطی مثبت آن‌ها می‌توان همه‌ی حالت‌های شار برای یک شبکه را تولید کرد. اما برای تولید همه‌ی فضای شار ممکن در یک شبکه‌ی متابولیکی، همه‌ی مدهای پایه لازم نیستند. بلکه زیرمجموعه‌ای از مدهای پایه به نام مدهای مولد شار<sup>۱</sup> که تعداد آن‌ها معمولاً خیلی کمتر از مدهای پایه‌ی شار است، برای این مقصود کافی هستند [۴۱]. این زیرمجموعه از مدهای پایه البته یکتا نیستند و بنابراین نمی‌توان مفهوم زیستی شفاف را به آن‌ها نسبت داد. به همین دلیل استفاده از این بردارها هرگز به اندازه‌ی مدهای پایه گسترده نبوده است [۲۰]. همچنین تعاریفی مانند رفتار کمینه‌ی متابولیکی<sup>۲</sup> و فضای برگشت پذیر متابولیکی<sup>۳</sup> نیز مفاهیمی جایگزین هستند که هدف آن‌ها توصیف یک فضای شار با حداقل تعداد بردارها است [۴۲]. با این وجود، مدهای پایه‌ی شار به دلیل انطباق آن‌ها بر مفهوم مسیر متابولیکی، بیشترین مقبولیت و کاربرد را در میان مفاهیم مشابه دارند.

### ۳-۱-۶ محاسبه‌ی مدهای پایه‌ی مولد برای یک زیرمجموعه‌ی خاص از فضای شار

استفاده از مدهای پایه‌ی شار محدود به فضای جواب کلی شار نیست. ممکن است مدهای پایه‌ای که یک زیرمجموعه‌ی خاص از فضای جواب را می‌سازند مد نظر باشد. برای مثال در روش تحلیل شار متابولیکی، یک مجموعه شار اندازه‌گیری شده‌ی تجربی وارد سیستم محدودیت‌های خطی می‌شوند. و باعث می‌شوند که یک شار یکتا برای واکنش‌های شبکه به دست بیاید. دانستن این که توزیع شار مورد نظر از چه مدهای پایه‌ای ساخته می‌شود گاهی مورد نظر است. بر این اساس روش‌هایی معرفی شده‌اند که یک

<sup>1</sup> Generating Flux Mode

<sup>2</sup> Minimal Metabolic Behaviour (MMB)

<sup>3</sup> Reversible Metabolic Space (RMS)

توزیع شار داده شده را به مدهای پایه‌ی مولد آن تجزیه می‌کنند. مزیت این روش‌ها این است که برای استفاده از آن‌ها دانستن مجموعه‌ی مدهای پایه از قبل لازم نیست و بنابراین به راحتی می‌توانند بر روی شبکه‌های مقیاس ژنوم اعمال گردند [۴۳، ۴۴].

همچنین روش تحلیل تعادل شار اغلب شامل یافتن یک توزیع شار بهینه براساس یک تابع هدف می‌باشد. توزیع شار شرایط بهینه معمولاً یکتا نیست و یک فضای جواب برای این شرایط به دست می‌آید. فضای شار ناشی از شرایط بهینه همانند فضای شار اصلی یک چند وجهی محدب بوده و دارای مدهای پایه است. در چند سال اخیر روش‌ها و ابزارهایی برای به دست آوردن این مدهای پایه به دست آمده است. ضمن این که نتایج اعمال این روش‌ها بر روی شبکه‌های مقیاس ژنوم نشان داده‌اند که تعداد زیاد مدهای پایه‌ی فضای جواب بهینه، ناشی از ترکیب مسیرهای مختلف در تعداد معدودی از زیرشبکه‌ها هستند و در سایر نقاط شبکه همه‌ی مدهای پایه از مسیر یکسانی می‌گذرند. با ترکیب حالات مسیرها تنوع کاذبی از مدهای پایه به وجود می‌آید که منجر به تعداد زیاد آن‌ها می‌گردد. بنابراین روشی ارایه شده که توسط آن می‌توان مدهای پایه‌ی این فضای جواب را به صورت فشرده‌تری نمایش داد. به این صورت که به جای نمایش همه‌ی مدهای پایه‌ی شبکه، مدهای پایه‌ی زیرشبکه‌های گفته شده به دست می‌آید و از هر ترکیب آن‌ها می‌توان مدهای پایه‌ی شبکه‌ی اصلی را ساخت [۴۵]. بر همین اساس مفهومی به نام ماژول شار<sup>۱</sup> مطرح شده است که زیرشبکه‌های مذکور را مشخص می‌کند. یک ماژول شار به ازای یک فضای شار داده شده تعریف می‌گردد و عبارت است از زیرشبکه‌ای از یک شبکه‌ی متابولیکی که به ازای همه‌ی نقاط فضای جواب داده شده، شار گذرنده از مرزهای آن ثابت است. نحوه‌ی به دست آوردن ماژول‌های شار به این صورت است که با استفاده از تحلیل تغییرپذیری شار، مجموعه‌ی واکنش‌هایی که همواره دارای شار ثابتی هستند مشخص می‌شود. سپس این واکنش‌ها از شبکه حذف می‌شوند. زیرشبکه‌هایی از شبکه که در اثر این عملیات به صورت ایزوله در می‌آیند ماژول‌های شار را مشخص می‌کنند. زیرا این زیرشبکه‌ها محصور در واکنش‌های با شار ثابت بوده‌اند و بنابراین پس از حذف واکنش‌های با شار ثابت، ارتباط آن‌ها با بقیه‌ی شبکه قطع می‌شود [۴۶]. با به دست آوردن ماژول‌های شار برای فضای جواب بهینه‌ی شبکه‌های متابولیکی مقیاس ژنوم، می‌توان به جای به دست آوردن

<sup>۱</sup> Flux Module



مدهای پایه برای کل شبکه، مدهای پایه را برای ماژول‌های شار به دست آورد. هر ترکیب مدهای پایه‌ی ماژول‌های شار یک مدشار برای شبکه‌ی کلی را معرفی می‌کند. با توجه به سهولتی که تعریف ماژول‌های شار برای فضای جواب بهینه‌ی شبکه‌های متابولیکی ایجاد می‌کند، سعی شده که این روش برای فضای جواب کلی حالت پایا نیز مورد استفاده قرار گیرد. بررسی‌ها نشان داده‌اند که در مورد فضای جواب کلی ماژول‌های شار مشاهده نمی‌شوند. بنابراین این روش برای فضای جواب کلی کارایی ندارد [۴۶]. البته این تلاش به تعریف مفهومی به نام  $k$ -module انجامیده است.  $k$ -module ها زیرشبکه‌هایی هستند که شار گذرنده از مرزهای آن‌ها ثابت نیست اما می‌تواند با استفاده از یک فضای  $k$  بعدی (با استفاده از  $k$  متغیر مستقل) توصیف شود. هرچه عدد  $k$  بتواند کوچکتر باشد به معنی این است که این شبکه دارای ارتباط کمتری با بقیه‌ی شبکه است. البته تاکنون روش مشخصی برای یافتن مدهای پایه از طریق تعریف  $k$ -module ها ارائه نشده است.

در این بخش به تاریخچه‌ی ابزارها و روش‌های محاسبه‌ی مدهای پایه‌ی شار پرداختیم. در بخش بعد به جنبه‌ی متفاوتی از شبکه‌های متابولیکی یعنی ماژولار بودن و رویکردهایی که برای تعریف ماژول در شبکه‌های متابولیکی معرفی شده است می‌پردازیم.

## ۲-۳ تجزیه‌ی شبکه‌های متابولیکی

تجزیه‌ی شبکه‌های متابولیکی به قسمت‌های کوچک‌تر همانند اکثر شبکه‌های بزرگ دیگر، راهکار طبیعی برای رویارویی با پیچیدگی موجود در آن‌ها است. از گذشته به منظور مطالعه‌ی متابولیسم، با استفاده از شهود، آن را به بخش‌هایی از قبیل گلیکولیز و تری کربوکسیلیک اسید تقسیم کرده‌اند. هم‌اکنون با فراهم شدن مدل‌های مقیاس ژنوم متابولیسم بسیاری بر لزوم تشخیص ماژول‌های شبکه به صورت بی‌طرفانه و بدون دخالت شهود انسان تاکید کرده‌اند، با این ایده که این رویکرد به کشف عملکردهایی منجر می‌شود که از طریق روش شهودی قابل تشخیص نبوده‌اند [۴۷]. بر همین اساس تحقیقات بسیاری بر اساس معیارها مختلف سعی کرده‌اند که قسمت‌هایی از شبکه‌ی متابولیکی را به عنوان ماژول‌های عملیاتی معرفی کنند و از این طریق درک این سیستم‌ها را ساده‌تر سازند.

### ۳-۲-۱ بررسی ساختار اتصالات در شبکه‌های متابولیکی

به منظور به دست آوردن دیدگاهی از ساختار کلی شبکه‌های متابولیکی تحقیقاتی بر روی خواص توپولوژیکی این شبکه‌ها انجام شده است. برای مثال بررسی‌های انجام شده بر روی توزیع درجه‌ی رئوس در موجودات مختلف نشان داده است که این شبکه‌ها با شبکه‌های تصادفی متفاوت بوده و بسیار مشابه به ساختاری هستند که در شبکه‌های اجتماعی و وب جهانی مشاهده می‌شوند [۴۸]. این نوع شبکه‌ها که به شبکه‌های بی‌مقیاس<sup>۱</sup> معروفند، شبکه‌هایی هستند که برخلاف شبکه‌های تصادفی، همگن نبوده و رئوس با درجه‌ی بسیار بالا در آن‌ها مشاهده می‌شود. این مشاهدات به ویژگی‌های ذاتی این نوع شبکه‌ها که قوام و تحمل‌پذیری خطا<sup>۲</sup> است نسبت داده شده است [۴۸]. صرف نظر از دلیل چنین مشاهداتی، وجود متابولیت‌هایی در شبکه‌های متابولیکی که در واکنش‌های بسیار زیادی حضور دارند، واقعیتی شناخته شده برای زیست‌شناسان است. از طرف دیگر بررسی‌های انجام گرفته بر روی ضریب خوشه‌ای<sup>۳</sup> بر وجود یک نوع حالت ماژولار در این شبکه‌ها دلالت دارد. در حالی که در شبکه‌های بی‌مقیاس معمولاً وجود گره‌های با اتصالات سراسری مانع از مشاهده‌ی حالت ماژولار می‌شوند. به منظور ارائه‌ی مدلی که پارامترهای مشاهده شده در شبکه‌های متابولیکی را توجیه می‌نماید، یک ساختار ماژولار سلسله‌مراتبی برای این شبکه‌ها پیشنهاد شده است که با بسیاری از اندازه‌گیری‌ها در این شبکه‌ها سازگار است. در این ساختار، ماژول‌ها وجود دارند اما همه‌ی آن‌ها به یک راس مرکزی وصل می‌شوند. سپس این ساختار به طور بازگشتی تکرار می‌شود [۴۹].

### ۳-۲-۲ تجزیه‌ی شبکه‌های متابولیکی به منظور یافتن ماژول‌های عملیاتی

در راستای رویکرد معرفی ماژول‌های مبتنی بر شبکه، یعنی ماژول‌هایی که فارغ از شهود زیست شناختی و بر اساس روش‌های مبتنی بر شبکه به دست می‌آیند، تحقیقات مختلفی بر اساس معیارهای مختلف از جمله ساختار ساختار سراسری و یا محلی اتصالات در گراف شبکه‌های متابولیکی، این

<sup>۱</sup> Scale-Free Networks

<sup>۲</sup> Error Tolerance

<sup>۳</sup> Clustering Coefficient

شبکه‌ها را به زیرشبکه‌هایی تجزیه کرده و این زیرشبکه‌ها را با ماژول‌های شناخته‌شده‌ی زیستی مقایسه نموده‌اند.

برای مثال شوستر با در نظر گرفتن این مطلب که در شبکه‌های متابولیکی متابولیت‌های با درجه‌ی بالا معمولاً در مرزهای زیرشبکه‌ها حضور دارند، با در نظر گرفتن یک حد آستانه‌ی ۴ برای درجه‌ی گره‌ها متابولیت‌هایی که دارای درجه‌ی راس بالایی هستند را از شبکه‌ی متابولیکی حذف نموده است. با اعمال این روش یک شبکه‌ی مقیاس ژنوم به ۱۳ زیرشبکه تقسیم شده که تا حدودی بر زیرشبکه‌های شناخته شده‌ی زیستی نیز انطباق داشته‌اند [۵۰].

در روشی دیگر مبتنی بر همین چارچوب با استفاده از خوشه‌بندی مارکوف<sup>۱</sup> زیرشبکه‌ها تشخیص داده می‌شوند. روش خوشه‌بندی مارکوف زیرشبکه‌ها را این‌گونه تعریف می‌کند که هنگامی که از داخل یک زیرشبکه شروع به حرکت کنیم، باید احتمال ماندن در داخل زیرشبکه از بیرون رفتن از آن بیشتر باشد [۵۱]. این روش به صورت یک برنامه ارائه گردیده و امکان تعیین دستی مرزهای زیرشبکه‌ها نیز به آن افزوده شده است [۵۲].

گروهی اطلاعات محلی شبکه، یعنی درجه‌ی راس‌ها را برای تجزیه‌ی شبکه کافی ندانسته و بر این اساس به تجزیه‌ی شبکه بر اساس اطلاعات سراسری پرداخته‌اند. در این روش با استفاده از روش‌های تجزیه‌ی موجود برای گراف‌های ساده، این گراف‌ها بر اساس اتصالات سراسری به زیرشبکه‌هایی تجزیه شده‌اند [۵۳].

در رویکردی دیگر بر این ایده تاکید شده است که در شبکه‌های پیچیده ساختار توپولوژیکی ماژول‌ها رابطه‌ی تنگاتنگی با کارکرد آن دارد. بر این مبنا با قیاس متابولیسم با شبکه‌های اجتماعی وجود نقش‌هایی برای مواد در شبکه‌ی متابولیکی پیشنهاد شده است. این نقش‌ها با تعریف پارامترهای درجه‌ی داخل ماژولی<sup>۲</sup> و ضریب شرکت‌کنندگی<sup>۳</sup> برای هر متابولیت مشخص می‌شود. همچنین معیاری به نام ماژولاریتی تعریف شده و تجزیه‌ی شبکه با هدف پیشینه کردن این معیار، با استفاده از الگوریتم

<sup>1</sup> Markov Clustering

<sup>2</sup> Within Module Degree

<sup>3</sup> Participating Coefficient

شبیه‌سازی سرد کردن فلزات<sup>۱</sup> انجام گرفته است. یکی از مشاهدات مهم ارائه شده در این تحقیق این است که در اکثر شبکه‌های متابولیکی، با انجام تجزیه‌ی گفته شده، ۸۰ درصد از متابولیت‌ها فقط با متابولیت‌های داخل ماژول خود ارتباط داشته‌اند [۵۴]. بنابراین تجزیه‌ی شبکه‌های متابولیکی بر اساس متابولیت‌ها کاملاً با ساختار ازتباطات در شبکه‌های متابولیکی هماهنگی دارد.

تعریف ماژول‌های شار و تعمیم آن‌ها به  $k$ -module ها نوعی تجزیه‌ی شبکه‌های متابولیکی به ماژول‌های عملیاتی می‌باشد. البته ماژول‌های شار برای یک حالت خاص زیستی تعریف می‌گردند (حالت بهینه‌ای که با استفاده از یک تابع هدف زیستی تعریف می‌گردد).  $k$ -module ها از طرف دیگر برای تعریف زیرشبکه‌هایی مفیدند که با توجه به ارتباط محدودی که با بقیه‌ی شبکه دارند، احتمالاً می‌توان فعالیت‌های متابولیکی به خصوصی را برای آن‌ها تصور کرد. ارتباط یک ماژول با بقیه‌ی شبکه در تعریف  $k$ -module ها به شکل خاصی بیان گردیده که با ماهیت مدل‌های مبتنی بر محدودیت هماهنگی دارد. زیرا این مفهوم با استفاده از متغیرها (شار واکنش‌ها) و وابستگی‌های خطی بین آن‌ها تعریف شده و بنابراین نتایج حاصل از آن‌ها می‌تواند در مدل مبتنی بر محدودیت اطلاعات مفیدی را دربر داشته باشد. در واقع اگر برای یک زیرشبکه، حداقل  $k$  ممکن را پیدا کنیم به طوری که آن شبکه بتواند یک  $k$ -module باشد، میزان ارتباط آن زیرشبکه با بقیه‌ی شبکه را می‌توان با معیار  $k$  سنجید [۵۵].

### ۳-۲-۳ تجزیه‌ی شبکه‌های متابولیکی و مدهای پایه

اولین تلاش برای استفاده از تجزیه‌ی شبکه برای محاسبه‌ی مدهای پایه را می‌توان تجزیه‌ی شوستر با استفاده از تعیین گره‌های با درجه‌ی بالا به عنوان گره مرزی دانست. در این روش پس از تعیین گره‌های مرزی و تعیین زیرشبکه‌ها، مسیرهای داخل زیرشبکه‌ها به طور مستقلی به دست می‌آید [۵۰]. البته توسط کالتا<sup>۲</sup> به این واقعیت اشاره شده است که به دست آوردن مدهای پایه‌ی یک زیرشبکه از شبکه‌ی متابولیکی می‌تواند گمراه کننده باشد. زیرا مدهای پایه‌ی یک زیرشبکه ممکن است بخشی از هیچ مد پایه‌ی شبکه‌ی کلی نباشند. بنابراین ممکن است مسیرهایی معرفی شوند که با واقعیات تجربی انطباق

<sup>۱</sup> Simulated Annealing

<sup>۲</sup> Kaleta

نداشته باشند. بر همین اساس مفهوم الگوهای پایه‌ی شار<sup>۱</sup> معرفی شده‌اند که آن‌ها مسیرهایی از واکنش‌های یک زیرشبکه را نشان می‌دهند که می‌توانند بخشی از یک مدپایه‌ی سراسری باشند [۵۶].

همچنین تعریف ماژول‌های شار به روشی برای محاسبه‌ی مدهای پایه‌ی مربوط به توزیع‌های بهینه شار منجر می‌شود. در این روش مدهای پایه می‌توانند به راحتی از ترکیب مدهای پایه‌ی ماژول‌های شار به دست بیایند. اما تعریف ماژول‌های شار به توزیع شار بهینه محدود می‌گردد و فضای جواب کلی توزیع‌های شار شبکه دارای ماژول‌های شار نمی‌باشد. بنابراین این روش نمی‌تواند برای به دست آوردن مدهای پایه‌ی شبکه در حالت کلی به کار رود. هرچند تعریف  $k$ -module ها از تعمیم ماژول‌های شار به فضای جواب کلی توزیع‌های شار شبکه به دست آمده است. اما رابطه‌ی مدهای پایه‌ی زیرشبکه‌های ناشی از  $k$ -module ها نامشخص است. بنابراین چارچوب و روشی که بتوان با استفاده از مدهای پایه‌ی زیرشبکه‌ها، مدهای پایه‌ی شبکه‌ی اصلی را محاسبه نمود تاکنون ارائه نشده است.

---

<sup>1</sup> Elementary Flux Pattern

۴

## فصل چهارم

### روش پیشنهادی

## روش پیشنهادی

در این فصل روش پیشنهادی برای محاسبه‌ی مدهای پایه از طریق تجزیه‌ی شبکه توضیح داده می‌شود. در این روش شبکه‌ی متابولیکی با تجزیه‌ی شبکه و تعریف درشت‌واکنش‌ها<sup>۱</sup> به یک شبکه‌ی درشت‌دانه‌تر<sup>۲</sup> تبدیل می‌شود. در شبکه‌ی حاصل، هریک از درشت‌واکنش‌ها نماینده‌ی یکی از مدهای پایه‌ی زیرشبکه‌ها هستند. این کار با این هدف انجام می‌گیرد که با درشت‌دانه‌سازی<sup>۳</sup> شبکه‌ی متابولیکی، اندازه‌ی مساله (یعنی تعداد متابولیت‌ها و واکنش‌ها) کاهش یابد. ضمن این که پس از درشت‌دانه‌سازی، مسیرهایی که عملکرد متابولیکی یکسانی دارند (مسیرهای افزونه)، به واکنش‌های یکسانی متناظر می‌گردند. در نتیجه یک نوع فشرده‌سازی انجام می‌گیرد.

از لحاظ شهودی درشت‌واکنش‌ها مفاهیم نا‌آشنایی نیستند. با توجه به این که هر مسیر متابولیکی نشان دهنده‌ی تبدیل متابولیت‌ها، البته از طریق یک سلسله واکنش، می‌باشد، بیان خلاصه شده‌ی چنین تبدیلیه وسیله‌ی یک واکنش دور از ذهن نمی‌باشد. در شکل ۴-۱ ایده‌ی تبدیل مسیرهای متابولیکی به واکنش‌های معادل نشان داده شده است.

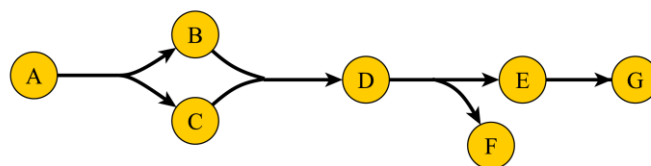
---

<sup>۱</sup> Coarse Reaction

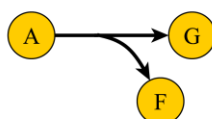
<sup>۲</sup> Coarse

<sup>۳</sup> Coarsening

یک مسیر متابولیکی



واکنش معادل



شکل ۴-۱ ایده‌ی معادل‌سازی یک مسیر متابولیکی با یک واکنش

با توجه به این که هریک از مدهای پایه، متناظر با تبدیل متابولیت‌های خارجی به یکدیگر، با یک نسبت خاص و ثابت هستند، به وضوح می‌توان تناظر بین یک مد پایه با یک واکنش مابین متابولیت‌های خارجی شبکه را تصور نمود. همچنین از آن جایی که مجموعه‌ی مدهای پایه همه‌ی تبدیلات ممکن از میان یک شبکه نشان می‌دهند، می‌توان تصور کرد که مجموعه‌ی درشت‌واکنش‌های بین متابولیت‌های خارجی می‌توانند به عنوان خلاصه‌ی از یک شبکه‌ی متابولیکی به کار روند.

در این تحقیق، با انتخاب زیرشبکه‌های مجزا، و خلاصه‌سازی آن‌ها با روش گفته شده، شبکه به شبکه‌ای درشت‌دانه تبدیل می‌گردد. اگر انتخاب زیرشبکه‌ها به گونه‌ای باشد که به صورت کلی تعداد مسیرهای گذرنده از آن‌ها از تعداد واکنش‌های موجود در آن‌ها کمتر باشد، تبدیل مسیرهای زیرشبکه‌ها به واکنش‌های معادل، با کاهش تعداد واکنش‌های شبکه همراه خواهد بود. به این ترتیب، شبکه به شبکه‌ی فشرده‌تری با اندازه‌ی کوچکتر تبدیل خواهد شد. با محاسبه‌ی مدهای پایه‌ی شبکه‌ی درشت‌دانه می‌توان آن‌ها را به مدهای پایه‌ی شبکه‌ی اصلی تبدیل کرد.

از دیدگاه دیگر می‌توان گفت که بررسی شبکه‌ی درشت‌دانه و به طور خاص محاسبه‌ی مدهای پایه‌ی آن، مشخص می‌سازد که مدهای پایه‌ی زیرشبکه‌ها چگونه می‌توانند به یکدیگر اتصال یابند تا مدهای پایه‌ی شبکه‌ی کلی ساخته شود. بنابراین با استفاده از این روش ارتباط بین مدهای پایه‌ی زیرشبکه‌ها و مدهای پایه‌ی شبکه‌ی کلی برقرار می‌گردد.



قبل از بیان نحوه‌ی درشت‌دانه‌سازی<sup>۱</sup> شبکه‌های متابولیکی باید ابتدا چارچوبی برای تجزیه‌ی این شبکه‌ها به زیرشبکه‌های کاملاً مجزا به دست آید. این کار در بخش ۴-۱ انجام گرفته است.

## ۴-۱ به دست آوردن چارچوبی برای تجزیه‌ی شبکه‌های متابولیکی به زیرشبکه‌های جدا از هم

برای این که فرایند تجزیه و درشت‌دانه‌سازی شبکه بدون هیچ‌گونه ابهامی صورت گیرد، چارچوب مشخصی برای این عملیات اتخاذ می‌نماییم. هریک از مفاهیم زیرشبکه‌ها و مدهای زیرشبکه‌ها تا حدودی مفاهیمی غیردقیق هستند که باید به روشنی تعریف گردند. همچنین برای این که این تعاریف به درستی معنا پیدا کنند و نیز هدف نهایی، یعنی اتصال مدهای پایه‌ی زیرشبکه‌ها و تولید مدهای پایه‌ی شبکه‌ی کلی، تحقق پیدا کند، شرایطی را برای نحوه‌ی تجزیه‌ی شبکه قرارداد می‌کنیم.

### ۴-۱-۱ نحوه‌ی تعریف زیرشبکه‌ها

نحوه‌ی تعریف زیرشبکه‌ها در روش‌های مختلف به شکل‌های متفاوتی مطرح شده است. برای مثال در روش شوستر ابتدا گره‌های با درجه‌ی بالا مشخص می‌شوند. سپس هر مجموعه‌ای از متابولیت‌ها و واکنش‌ها که محصور در چنین مجموعه متابولیت با درجه‌ی بالا هستند یک زیرشبکه را تعریف می‌کنند. یعنی در واقع متابولیت‌های با درجه‌ی بالا مرز زیر شبکه‌ها را مشخص می‌سازند [۵۰]. همچنین رویکرد دیگر می‌تواند این باشد که هر زیر مجموعه‌ی دلخواهی از واکنش‌ها به عنوان یک زیرشبکه در نظر گرفته شود [۴۶، ۵۵، ۵۷].

در هر صورت اگر بخواهیم زیرشبکه‌ها را به گونه‌ای تعریف کنیم که مدهای پایه برای آن‌ها قابل تعریف و محاسبه باشد باید بتوان برای آن زیرشبکه، مدلی شامل متغیرها (واکنش‌ها) و معادلات خطی

<sup>1</sup> Coarsening

(متابولیت‌ها) را تشخیص داد. به طوری که متغیرها و قید<sup>۱</sup>های مدل زیرشبکه، زیر مجموعه‌ای از متغیرها و قیدهای مدل مربوط به شبکه‌ی اصلی باشد.

مدل مبتنی بر محدودیت شامل اجزای زیر است:

۱. متابولیت‌های داخلی

۲. متابولیت‌های خارجی (البته متابولیت‌های خارجی نقشی در قیدهای شبکه و محاسبه‌ی مدهای پایه ندارند)

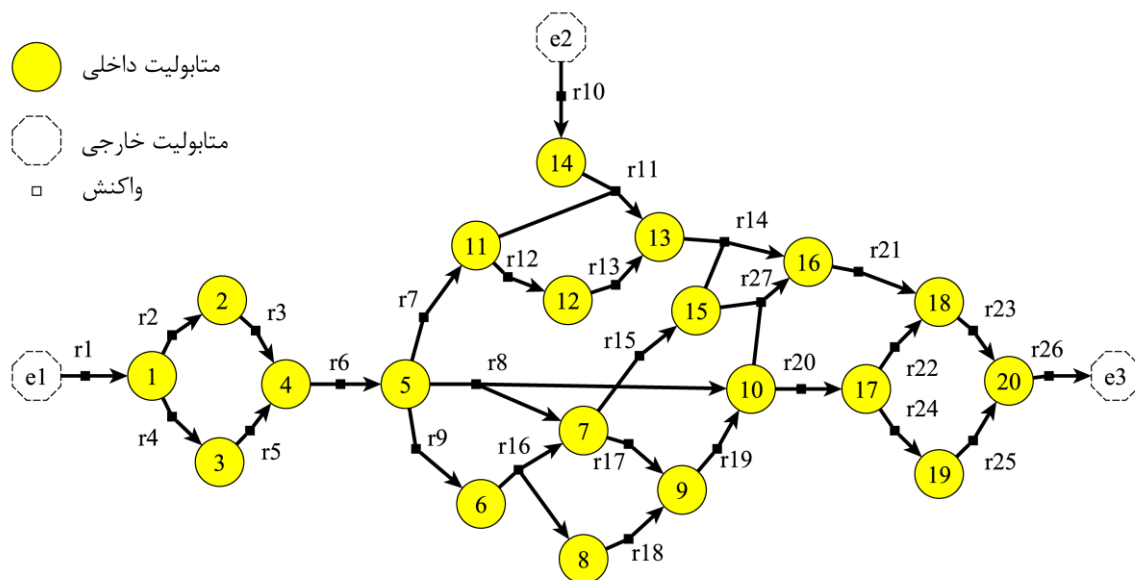
۳. واکنش‌های موجود بین متابولیت‌ها

برای مثال در شبکه‌ی نشان داده شده در شکل ۲-۴ مجموعه‌ی واکنش‌ها (R)، مجموعه‌ی متابولیت‌های خارجی (E) و مجموعه‌ی متابولیت‌های داخلی (M) عبارتند از:

$$R = \{r1, r2, \dots, r26\}$$

$$M = \{m1, m2, \dots, m21\}$$

$$E = \{e1, e2, e3\}$$



شکل ۲-۴ یک شبکه‌ی متابولیکی نمونه

<sup>1</sup> Constraint

برای این که بتوان زیر شبکه‌ای مشخص کرد که مدهای پایه برای آن قابل تعریف باشند، باید قاعداً زیر شبکه‌ای مشخص شده، اجزای یک مدل مبتنی بر محدودیت (متابولیت‌های داخلی و خارجی و واکنش‌ها) را برای ما مشخص سازد.

یک روش این است که زیر شبکه را با انتخاب متابولیت‌های آن مشخص کرد. آن گاه متابولیت‌های انتخاب شده را با همه‌ی واکنش‌های متصل به آن‌ها به عنوان یک زیر شبکه در نظر گرفت. از میان متابولیت‌های باقی مانده‌ی شبکه‌ی اصلی، آن‌هایی که با واکنش‌های در نظر گرفته شده اتصال دارند به عنوان متابولیت‌های خارجی برای این زیر شبکه در نظر گرفته می‌شوند.

برای مثال در مورد شبکه‌ی متابولیکی شکل ۴-۲ با انتخاب متابولیت‌های ۶، ۷، ۸ و ۹، زیر شبکه‌ی متناظر به صورت زیر تعریف می‌شود.

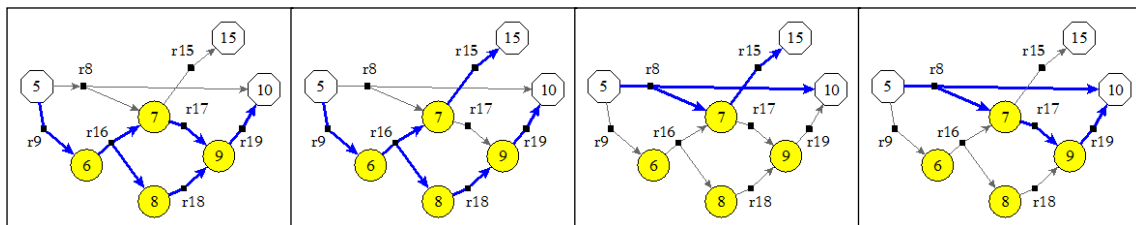
$$N_1 :$$

$$M_1 = \{6, 7, 8, 9\}$$

$$R_1 = \{r8, r9, r15, r16, r17, r18, r19\}$$

$$E_1 = \{5, 10, 15\}$$

برای زیر شبکه‌ی  $N_1$  می‌توان مدهای پایه را به صورت شکل ۴-۳ به دست آورد



شکل ۴-۳ مدهای پایه برای زیر شبکه‌ی متناظر با مجموعه‌ی متابولیت‌های ۶، ۷، ۸ و ۹

به این ترتیب مدهای پایه‌ی به دست آمده را می‌توان به عنوان مسیرهایی از داخل زیر شبکه تعبیر کرد که توسط آن‌ها بقیه‌ی متابولیت‌های شبکه‌ی اصلی (به طور خاص آن‌هایی که به زیر شبکه متصل هستند) می‌توانند به یکدیگر تبدیل شوند.

نحوه‌ی تعریف یک زیر شبکه با استفاده از انتخاب یک زیر مجموعه از متابولیت‌ها

اگر شبکه‌ی  $N$  یک شبکه‌ی متابولیکی باشد که مجموعه‌ی متابولیت‌های آن  $M$  و مجموعه‌ی واکنش‌های آن  $R$  باشند، با انتخاب یک مجموعه‌ی  $M_I$  (به‌طوری که  $M_I \subseteq M$ )، می‌توان زیرشبکه‌ی  $N_I$  حاصل از این انتخاب را به صورت زیر مشخص نمود:

۱. مجموعه‌ی متابولیت‌های داخلی برابر مجموعه‌ی  $M_I$  است.

۲. مجموعه‌ی واکنش‌های  $N_I$  برابر مجموعه‌ی واکنش‌هایی در  $R$  است که به متابولیت‌های مجموعه‌ی  $M_I$  متصل هستند.

۳. مجموعه‌ی متابولیت‌های خارجی ( $E_I$ ) برابر متابولیت‌هایی از مجموعه‌ی  $M - M_I$  است که به متابولیت‌های مجموعه‌ی  $M$  از طریق حداقل یک واکنش متصل هستند.

#### ۴-۱-۲ نحوه‌ی تجزیه‌ی شبکه

منظور از تجزیه‌ی شبکه تقسیم آن به تعدادی زیرشبکه است. هنگامی که قصد تجزیه‌ی شبکه به زیرشبکه‌ها و بررسی برهم‌کنش زیرشبکه‌های متعدد را داریم یکی از مسایلی که مطرح می‌شود امکان همپوشانی بین زیرشبکه‌هاست. همپوشانی در اینجا به معنی وجود واکنش‌ها یا متابولیت‌های مشترک بین زیر شبکه‌های مختلف است. برای این که مطمئن شویم که وجود متابولیت‌ها یا واکنش‌های مشترک در بین زیرشبکه‌ها، باعث ایجاد وابستگی بین زیرشبکه‌های مختلف نخواهد شد، تجزیه را به نحوی انجام می‌دهیم که همه‌ی زیرشبکه‌ها از یکدیگر کاملاً جدا بوده و هر متابولیت و واکنشی متعلق به تنها یکی از زیر شبکه باشد.

شرط جدایی زیرشبکه‌ها حتماً در مورد شبکه‌های متابولیکی محدود کننده خواهد بود. زیرا شبکه‌های متابولیکی را معمولاً نمی‌توان به بخش‌های کاملاً مجزا تقسیم کرد. بنابراین گروهی از متابولیت‌ها به نام متابولیت‌های میانی به همراه واکنش‌هایی که فقط به این متابولیت‌ها متصلند از عملیات تجزیه‌ی شبکه مستثنی می‌شوند. این متابولیت‌ها به جای این که به یکی از زیرشبکه‌ها اختصاص داده شوند، نقش رابط بین زیرشبکه‌ها را ایفا می‌کنند. یعنی هنگام بررسی داخلی زیرشبکه‌ها این متابولیت‌ها به عنوان متابولیت‌های خارجی در نظر گرفته می‌شوند و هنگام بررسی رابطه‌ی بین زیرشبکه‌ها به عنوان متابولیت داخلی نقش اتصال زیرشبکه‌ها را بر عهده دارند.

بنابراین از نظر ما حاصل عملیات تجزیه معادل تقسیم یک شبکه‌ی متابولیکی به بخش‌های زیر است.

۱. زیرشبکه‌ها: تعدادی زیرشبکه که به طور کامل از یکدیگر جدا هستند. یعنی هیچ متابولیت یا واکنشی بین آن‌ها مشترک نیست. این زیرشبکه‌ها با تعریف بخش ۴-۱ از زیرشبکه‌ها سازگاری دارند. یعنی این زیرشبکه‌ها شامل تعدادی متابولیت و همه‌ی واکنش‌های متصل به آن‌ها هستند.

۲. شبکه‌ی مرکزی: متابولیت‌های میانی، متابولیت‌های خارجی و واکنش‌هایی که تنها به متابولیت‌های میانی یا متابولیت‌های خارجی متصل هستند.

### تجزیه‌ی واجد شرایط شبکه

اگر برای شبکه‌ی متابولیکی مجموعه‌ی متابولیت‌های داخلی برابر  $M$ ، مجموعه‌ی واکنش‌ها برابر  $R$ ، و مجموعه‌ی متابولیت‌های خارجی برابر  $E$  باشد، آن‌گاه یک تجزیه‌ی واجد شرایط برای این شبکه شامل تعدادی زیرشبکه  $N_1 \dots N_k$  به همراه یک شبکه‌ی مرکزی  $N_C$  با شرایط زیر خواهد بود.

۱. هریک از زیر شبکه‌ی  $N_i$  دارای یک مجموعه‌ی  $M_i$  از متابولیت‌های داخلی است. مجموعه‌ی واکنش‌های آن  $(R_i)$  برابر مجموعه‌ی همه‌ی واکنش‌های موجود در  $R$  است که حداقل به یک متابولیت در  $M_i$  متصل هستند. مجموعه‌ی متابولیت‌های خارجی  $(E_i)$  برای هریک از این زیرشبکه‌ها برابر مجموعه‌ای از متابولیت‌های موجود در  $(M \cup E) - M_i$  است که حداقل به یکی از واکنش‌های  $R_i$  متصل هستند.

$$M_i \subseteq M$$

$$R_i = \{r \in R : \exists m \in M_i : r \text{ is connected to } m\}$$

$$E_i = \{x \in (M \cup E) - M_i : \exists r \in R_i : x \text{ is connected to } r\}$$

۲. شبکه‌ی مرکزی دارای مجموعه‌ی متابولیت‌های  $M_C$  است. مجموعه‌ی واکنش‌های شبکه‌ی مرکزی برابر مجموعه‌ی واکنش‌هایی در  $R$  است که تنها به متابولیت‌های  $M_C$  یا  $E$  متصل هستند. مجموعه‌ی متابولیت‌های خارجی شبکه‌ی مرکزی  $(E_i)$  برابر مجموعه‌ی متابولیت‌های خارجی شبکه‌ی اصلی ( $E$ ) است.

$$M_C \subseteq M$$

$$R_C = \{r \in R : \nexists m \in M - M_C : r \text{ is connected to } m\}$$

$$E_C = E$$

۳. بین متابولیت‌های داخلی زیرشبکه‌ها ( $M_i$ ) هیچ اشتراکی وجود ندارد. اما بین متابولیت‌های خارجی ( $E_i$ ) می‌تواند اشتراک وجود داشته باشد.

$$\forall i, j \in 1 \dots k, i \neq j: M_i \cap M_j = \emptyset$$

۴. بین مجموعه‌ی واکنش‌های زیرشبکه‌ها ( $R_i$ ) هیچ اشتراکی وجود ندارد.

$$\forall i, j \in 1 \dots k, i \neq j: R_i \cap R_j = \emptyset$$

۵. اجتماع مجموعه‌ی متابولیت‌های زیرشبکه‌ها به همراه مجموعه‌ی متابولیت‌های میانی برابر مجموعه‌ی متابولیت‌های شبکه‌ی اصلی است. همچنین اجتماع واکنش‌های زیرشبکه‌ها با واکنش‌های شبکه‌ی مرکزی برابر مجموعه‌ی واکنش‌های شبکه‌ی اصلی است.

$$\left( \bigcup_{i \in 1 \dots k} M_i \right) \cup M_C = M$$

$$\left( \bigcup_{i \in 1 \dots k} R_i \right) \cup R_C = R$$

با استفاده از یک افراز از متابولیت‌های شبکه با رعایت شرایطی می‌توان یک تجزیه‌ی شبکه به دست آورد. به این صورت که متابولیت‌ها به چند زیرمجموعه به همراه یک مجموعه متابولیت‌های میانی افراز می‌شوند. این افراز با این شرط انجام می‌گردد که هیچ واکنشی بین متابولیت‌های دو زیر مجموعه‌ی مختلف وجود نداشته باشد. اما متابولیت‌های یک زیر شبکه می‌توانند با متابولیت‌های میانی در ارتباط باشند. پس از این که چنین افرازی به دست آمد، برای به دست آوردن تجزیه‌ی متناظر کافی است برای هر یک از زیرمجموعه‌های حاصل از افراز (به غیر از زیرمجموعه‌ی متابولیت‌های میانی) زیرشبکه‌ی متناظر را با اضافه کردن همه‌ی واکنش‌های متصل به آن‌ها به دست بیاوریم. برای به دست آوردن شبکه‌ی مرکزی، واکنش‌هایی را که فقط به متابولیت‌های میانی متصل هستند به زیرشبکه‌ی مربوط به آن‌ها اختصاص می‌دهیم.

### افراز واجد شرایط برای متابولیت‌ها

برای یک شبکه‌ی متابولیکی  $N$  با مجموعه‌ی متابولیت‌های  $M$  و مجموعه‌ی واکنش‌های  $R$ ، یک افراز واجد شرایط متابولیت‌ها شامل زیرمجموعه‌های  $M_1 \dots M_k$  به همراه زیرمجموعه‌ی  $M_C$  از متابولیت‌ها با شرایط زیر است.

۱. بین زیر مجموعه‌ها هیچ اشتراکی وجود ندارد و اجتماع این زیرمجموعه‌ها برابر مجموعه‌ی متابولیت‌ها است.

۲. بین متابولیت‌های دو زیر مجموعه‌ی متفاوت هیچ واکنشی وجود ندارد. اما بین متابولیت‌های هر زیرشبکه و متابولیت‌های میانی می‌تواند واکنش وجود داشته باشد.

### به دست آوردن یک تجزیه‌ی واجد شرایط از یک افراز واجد شرایط متابولیت‌ها

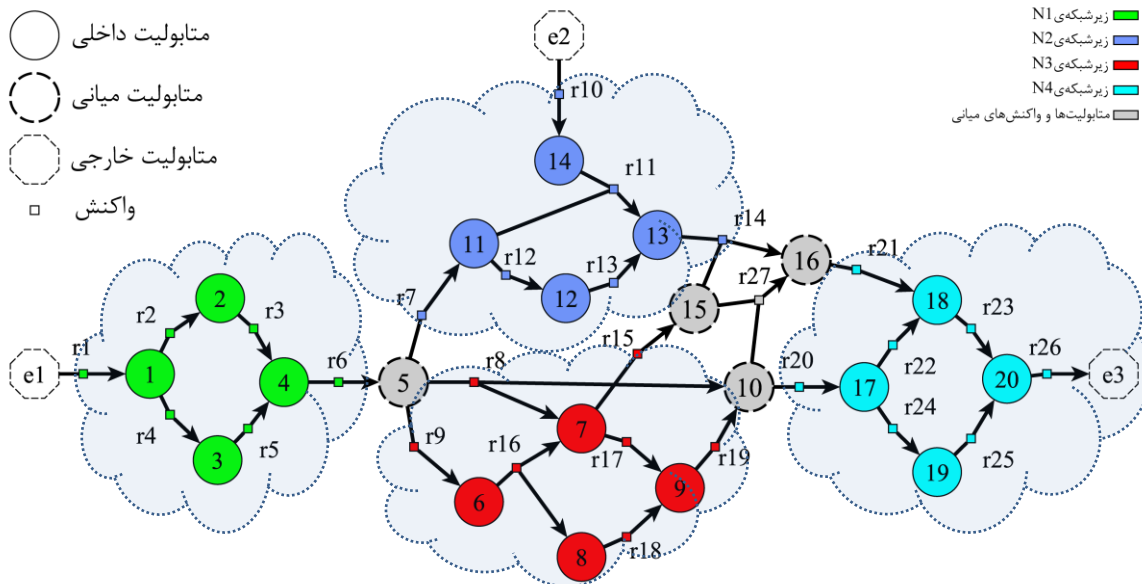
اگر مجموعه‌های  $M_1, \dots, M_k$  و  $M_C$  مجموعه‌های حاصل از یک افراز واجد شرایط برای متابولیت‌های شبکه‌ی  $N$  باشند. آن‌گاه یک تجزیه‌ی واجد شرایط برای  $N$  از راه زیر به دست می‌آید:

۱. به ازای هریک  $M_i$  ها یک زیرشبکه‌ی  $N_i$  تعریف می‌شود. مجموعه‌ی متابولیت‌های این زیرشبکه برابر  $M_i$  و مجموعه‌ی واکنش‌های آن برابر همه‌ی واکنش‌های موجود در  $N$  است که به متابولیت‌های موجود در  $M_i$  متصل هستند.

۲. یک زیرشبکه‌ی میانی  $N_C$  تعریف می‌شود که مجموعه‌ی متابولیت‌های آن برابر  $M_C$  و مجموعه‌ی واکنش‌های آن برابر مجموعه‌ای از واکنش‌های شبکه‌ی  $N$  است که تنها به متابولیت‌های  $M_C$  متصلند.

**مثال)** در مورد شبکه‌ی متابولیکی شکل ۴-۲ اگر مجموعه‌ی متابولیت‌ها را به زیرمجموعه‌های  $M_1, \dots, M_4$  و  $M_C$  افراز کنیم، که مجموعه‌ی  $M_C$  مجموعه‌ی متابولیت‌های میانی است، زیرشبکه‌های زیر به دست خواهد آمد که این زیرشبکه‌ها در شکل ۴-۴ مشخص شده‌اند:

|   |   |
|---|---|
| $N_1 :$                                     | $N_2 :$                                       |
| $M_1 = \{1, 2, 3, 4\}$                      | $M_2 = \{11, 12, 13, 14\}$                    |
| $R_1 = \{r1, r2, r3, r4, r5, r6\}$          | $R_2 = \{r7, r10, r11, r12, r13, r14\}$       |
| $E_1 = \{e1, 5\}$                           | $E_2 = \{5, 15, 16, e2\}$                     |
| $N_3 :$                                     | $N_4 :$                                       |
| $M_3 = \{6, 7, 8, 9\}$                      | $M_4 = \{17, 18, 19, 20\}$                    |
| $R_3 = \{r8, r9, r15, r16, r17, r18, r19\}$ | $R_4 = \{r20, r21, r22, r23, r24, r25, r26\}$ |
| $E_3 = \{5, 10, 15\}$                       | $E_4 = \{10, 16, e3\}$                        |

$N_C :$ 
 $M_C = \{5, 10, 15, 16\}$ 
 $R_C = \{r_{27}\}$ 
 $E_C = \{e_1, e_2, e_3\}$ 


شکل ۴-۴ تجزیه‌ی به دست آمده از افراز متابولیت‌ها برای شبکه‌ی شکل ۲-۴

## ۴-۲ به دست آوردن مدهای پایه از طریق درشت‌دانه سازی

نتیجه‌ی فرایند تجزیه‌ی شبکه که در بخش ۴-۱-۲ معرفی شد، یک تجزیه‌ی واجد شرایط برای شبکه است. چنین تجزیه‌ای شامل مجموعه‌ای از زیرشبکه‌ها به همراه شبکه‌ی مرکزی است. با داشتن یک تجزیه‌ی واجد شرایط برای شبکه می‌توان عملیات درشت‌دانه‌سازی را روی آن انجام داد. تجزیه‌ی واجد شرایط دارای این ویژگی است که در آن زیرشبکه‌ها با یکدیگر هیچ اشتراکی ندارند. بنابراین برای هر یک از این زیرشبکه‌ها می‌توان مدهای پایه را مستقلاً به دست آورد. هنگام به دست آوردن مدهای پایه در هر یک از زیرشبکه‌ها متابولیت‌های میانی به عنوان متابولیت‌های خارجی در نظر گرفته می‌شوند. با این فرض هر یک از مدهای پایه‌ی به دست آمده برای یک زیرشبکه به معنای یک مسیر تبدیل برای متابولیت‌های میانی به یکدیگر از طریق آن زیرشبکه هستند. این مسیرها بدون در نظر گرفتن بقیه‌ی شبکه است و تنها قیدهای ایجاد شده در داخل این زیرشبکه بر روی مرزها را نشان می‌دهند.



برای این که بدانیم این مدهای پایه به چه صورتی می‌توانند به یکدیگر متصل شوند تا مدهای پایه‌ی شبکه‌ی کلی را بسازند، از روش درشت‌دانه سازی شبکه استفاده می‌کنیم. به این صورت که ارتباطات بین مدهای پایه‌ی زیرشبکه‌ها در شبکه‌ی درشت‌دانه بررسی می‌شود. شبکه‌ی درشت‌دانه شبکه‌ای است که در آن به جای هر زیر شبکه، خلاصه‌ای از آن قرار داده می‌شود. صورت خلاصه شده‌ی هر زیرشبکه در واقع واکنش‌های مستقیمی بین متابولیت‌های میانی است که هریک معادل یک مد پایه در آن زیرشبکه می‌باشند. به این واکنش‌های مستقیم درشت‌واکنش می‌گوییم.

به دست آوردن شبکه‌ی درشت‌دانه شامل مراحل زیر است:

۱. ابتدا شبکه‌ی مرکزی که از فرایند تجزیه به دست آمده را در نظر می‌گیریم. این شبکه از میان واکنش‌ها و متابولیت‌های شبکه‌ی اصلی شامل متابولیت‌های میانی (به عنوان متابولیت‌های داخلی)، متابولیت‌های خارجی (به عنوان متابولیت‌های خارجی) و واکنش‌هایی که بین این متابولیت‌ها وجود دارند می‌باشد.

۲. برای هریک از زیرشبکه‌ها مدهای پایه به دست می‌آیند. سپس هریک از این مدهای پایه به درشت‌واکنشی بین متابولیت‌های میانی تبدیل می‌شوند. هریک از درشت‌واکنش‌های به‌دست‌آمده به شبکه‌ی مرکزی اضافه می‌شوند.

۳. پس از اضافه کردن همه‌ی درشت‌واکنش‌های مربوط به همه‌ی زیرشبکه‌ها، شبکه‌ی درشت‌دانه حاصل می‌شود.

در شبکه‌ی درشت‌دانه، اتصال درشت‌واکنش‌ها (که مدهای پایه‌ی زیرشبکه‌ها بوده‌اند) از طریق واکنش‌های میانی برقرار می‌شود و به این ترتیب بررسی ارتباطات زیرشبکه‌ها و مدهای پایه‌ی آن‌ها ممکن می‌شود. با بررسی ساختار ارتباطات این درشت‌واکنش‌ها در شبکه‌ی درشت‌دانه می‌توانیم بدانیم که اتصال کدام مدهای پایه از لحاظ استوکیومتری مجاز است. ضمن این که برای دانستن این که چه اتصالی از مدهای پایه‌ی زیرشبکه‌ها به یک مدپایه منجر خواهد شد، باید سعی کنیم اتصال مورد نظر دارای حداقل درشت‌واکنش‌های ممکن باشد. بنابراین ما مدهای پایه‌ی شبکه‌ی درشت‌دانه (مدهای پایه‌ی درشت‌دانه) را به دست آورده و سپس آن را برحسب واکنش‌های شبکه‌ی اصلی می‌نویسیم. در ادامه هریک از این مراحل به همراه مثال توضیح داده خواهد شد.

## ۴-۲-۱ تبدیل مدهای پایه به واکنش‌های متناظر

هر مدشار در واقع یک نسبت خاص بین شار واکنش‌های شبکه را در حالت پایا بیان می‌کند. بر اثر چنین شاری از واکنش‌ها مقدار هریک از متابولیت‌های داخلی ثابت می‌ماند اما متابولیت‌های خارجی ممکن است تولید یا مصرف شوند. به ازای هر مدشار، تولید و مصرف متابولیت‌های خارجی با نسبتی خاص صورت می‌گیرد که می‌توان از ماتریس استوکیومتری و بردار شار مربوطه به دست آورد.

### به دست آوردن نرخ تولید هریک متابولیت‌های خارجی به ازای یک مد شار

اگر یک شبکه‌ی متابولیکی  $N$  دارای ماتریس استوکیومتری  $S_{m \times r}$  باشد، متناظر با متابولیت‌های خارجی نیز می‌توان یک ماتریس  $E_{e \times r}$  تعریف کرد که هریک از سطرهای آن معادل با یک متابولیت خارجی و هریک از ستون‌های آن متناظر با یکی از واکنش‌های شبکه است. این ماتریس ضرایب استوکیومتری هریک از واکنش‌های شبکه را برای هریک از متابولیت‌های خارجی نشان می‌دهد. برای به دست آوردن نرخ تولید هریک از متابولیت‌های خارجی که از بردار شار  $v$  ناشی می‌شود کافی است ماتریس  $E$  را در بردار ستونی  $v$  ضرب کنیم.

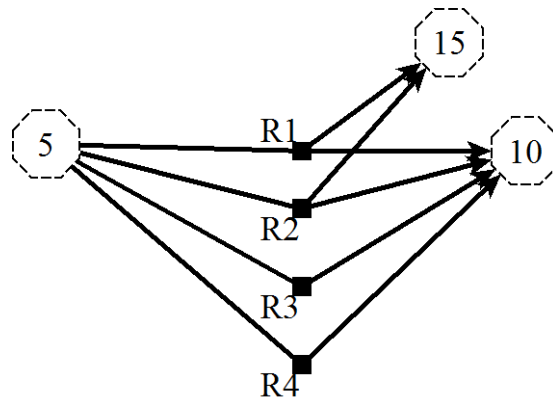
$$d = E.v$$

در این صورت هریک از عناصر بردار  $d$  مقدار نرخ تولید متابولیت خارجی متناظر را نشان می‌دهد. برای هر مدشار، نسبت بین نرخ‌های تولید متابولیت‌های خارجی یک نسبت ثابت است. یعنی منطبق با مفهوم یک واکنش، مواد را با نسبت مشخص به یکدیگر تبدیل می‌نماید. بنابراین هر مدشار و به‌طور ویژه هر مدپایه‌ی شار را می‌توان با یک واکنش بین متابولیت‌های خارجی متناظر کرد.

## ۴-۲-۲ ارتباط بین یک شبکه با شبکه‌ای متشکل از درشت‌واکنش‌های آن

با توجه به این که مدهای پایه‌ی یک زیرشبکه همه‌ی قابلیت‌های متابولیکی یک شبکه را توصیف می‌کنند، داشتن مجموعه‌ی مدهای پایه‌ی یک شبکه برای دانستن انواع حالت‌هایی که متابولیت‌های خارجی می‌توانند توسط این زیرشبکه به یکدیگر تبدیل شوند کافی است. هر توزیع شار شبکه را می‌توان به صورت یک ترکیب خطی از مدهای پایه‌ی شار به دست آورد. بنابراین می‌توان واکنش‌ها و متابولیت‌های یک شبکه را با درشت‌واکنش‌های آن جایگزین کرد. بدون این که از دیدگاه مرزهای این

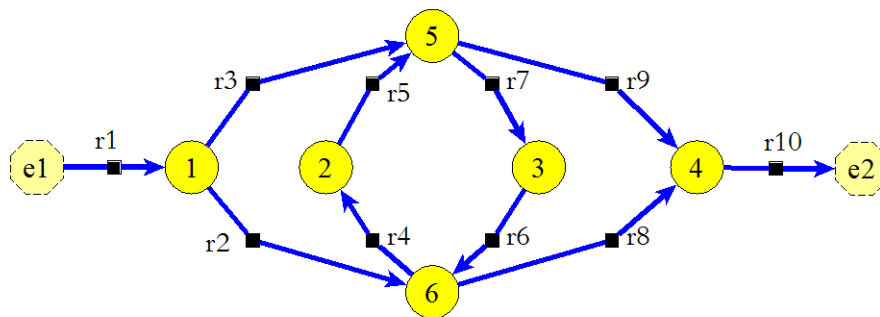
شبکه (متابولیت‌های خارجی) تغییری رخ دهد. برای مثال در مورد زیر شبکه‌ی  $N_3$  از شکل ۴-۴، شبکه‌ی متشکل از درشت‌واکنش‌ها به صورت شکل ۵-۴ خواهد بود.



شکل ۵-۴ شبکه‌ی متشکل از درشت‌واکنش‌ها برای زیر شبکه‌ی  $N_3$  در شکل ۴-۴

مساله‌ی دیگر نحوه‌ی تناظر بین فضای شار شبکه‌ی متشکل از درشت‌واکنش‌ها و شبکه‌ی اصلی است. هر توزیع شار یک شبکه را می‌توان به صورت یک ترکیب خطی مثبت از مدهای پایه‌ی آن نوشت. اما چنین توصیفی یکتا نیست و ممکن است بتوان توزیع شار را به چندین صورت به مدهای پایه‌ی سازنده‌ی آن تجزیه کرد. برای مثال توزیع شار نشان داده شده در شکل ۴-۶ را می‌توان به دو صورت به مدهای پایه تجزیه کرد

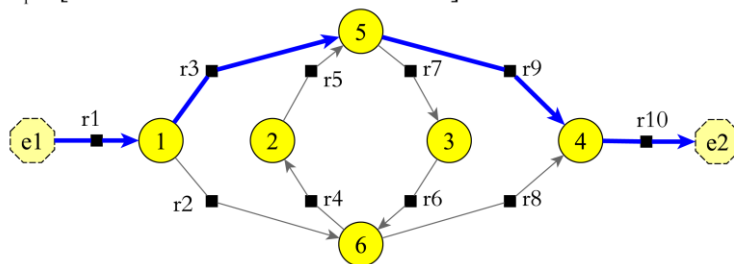
$$\begin{aligned}
 \mathbf{v} &= [v_1 \ v_2 \ v_3 \ v_4 \ v_5 \ v_6 \ v_7 \ v_8 \ v_9 \ v_{10}] \\
 &= [2 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 2] \\
 &= \mathbf{e}_3 + \mathbf{e}_4 \\
 &= [1 \ 0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 0 \ 1] \\
 &\quad + [1 \ 1 \ 0 \ 1 \ 1 \ 0 \ 0 \ 0 \ 1 \ 1] \\
 &= \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_5 \\
 &= [1 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 1] \\
 &\quad + [1 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1] \\
 &\quad + [0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 1 \ 0 \ 0 \ 0]
 \end{aligned}$$



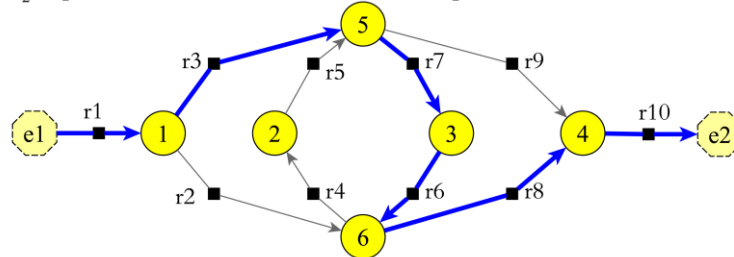
شکل ۴-۶ یک توزیع شار نمونه که به دو شکل قابل تجزیه است

در شکل ۴-۷ مدهای پایه‌ی  $\mathbf{e}_1 \dots \mathbf{e}_5$  نشان داده شده‌اند.

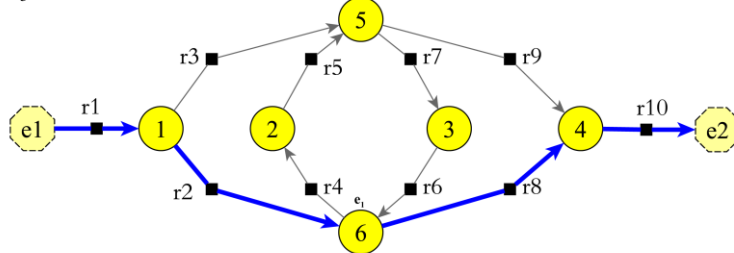
$$e_1 = [1 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 1]$$



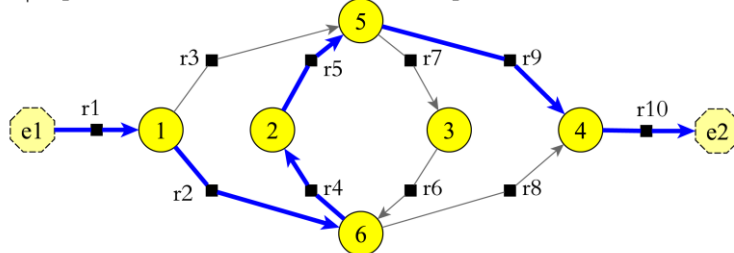
$$e_2 = [1 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1]$$



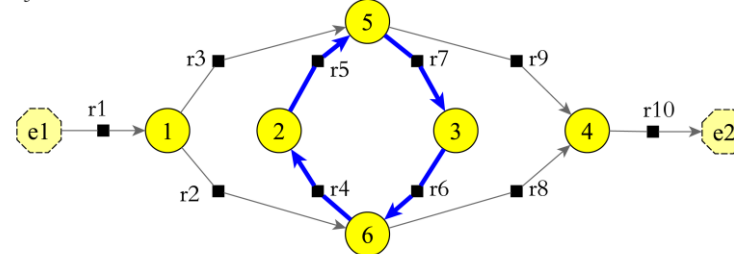
$$e_3 = [1 \ 0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 0 \ 1]$$



$$e_4 = [1 \ 1 \ 0 \ 1 \ 1 \ 0 \ 0 \ 0 \ 1 \ 1]$$



$$e_5 = [0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 1 \ 0 \ 0 \ 0]$$



شکل ۴-۷ مدهای پایه‌ی شبکه‌ی شکل ۴-۶

بنابراین می‌توان گفت که معادل با یک شار شبکه‌ی اصلی، ممکن است چندین شار در شبکه‌ی متشکل از درشت‌واکنش‌ها وجود داشته باشد.

در مقابل با داشتن یک توزیع شار برای شبکه‌ی درشت‌دانه این توزیع شار به طور یکتایی می‌تواند به شار در شبکه‌ی اصلی تبدیل شود. اگر از یک ضریب ثابت صرف نظر کنیم، هریک از مدهای پایه بیانگر توزیع شار مشخصی مانند  $\mathbf{e}$  در واکنش‌های شبکه می‌باشد. بنابراین شار  $a$  برای درشت‌واکنش متناظر با آن، شار  $a\mathbf{e}$  را برای شبکه‌ی اصلی در پی خواهد داشت. بنابراین اگر درشت‌واکنش‌های یک شبکه، یعنی  $R_1 \dots R_n$ ، به ترتیب دارای شار  $a_1 \dots a_n$  بوده و متناظر با مدهای پایه  $\mathbf{e}_1 \dots \mathbf{e}_n$  باشند، با جمع کردن اثر این درشت‌واکنش‌ها طبق رابطه‌ی ۱-۴، شار واکنش‌های شبکه به دست می‌آید.

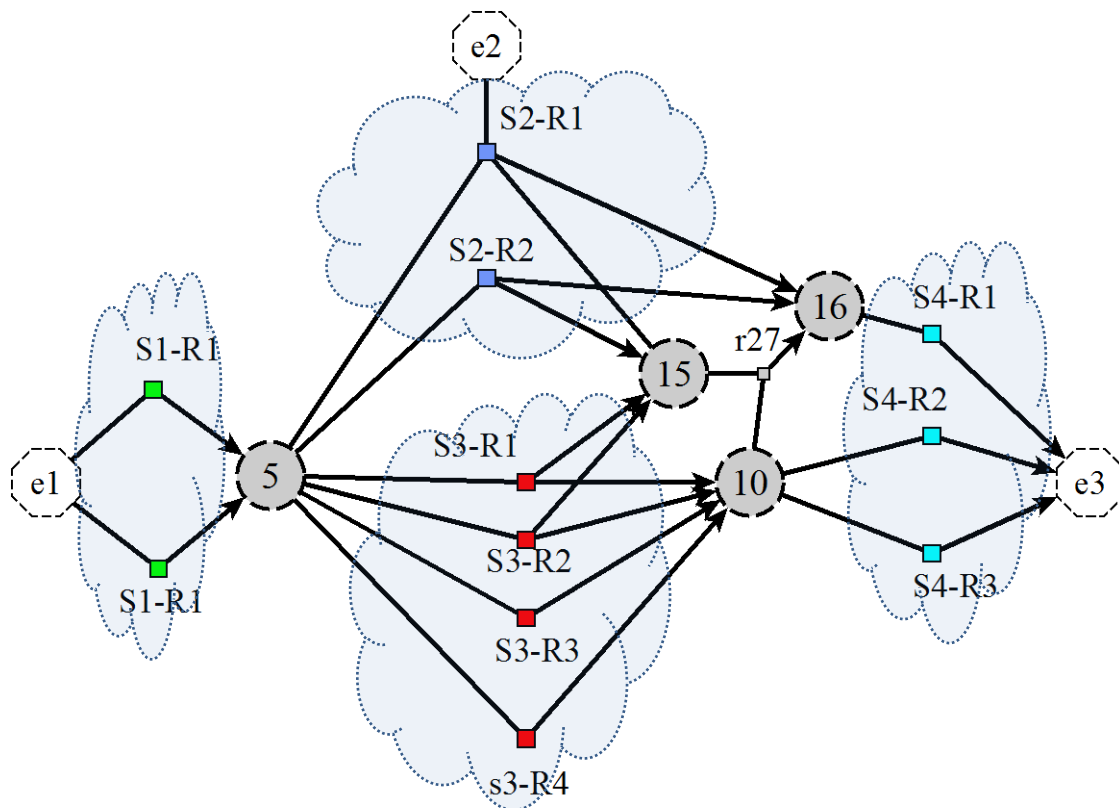
$$\sum_{i=1}^n a_i \mathbf{e}_i \quad 1-4$$

بنابراین هر توزیع شار بین درشت‌واکنش‌ها متناظر با دقیقاً یک توزیع شار واکنش‌های شبکه است اما هر توزیع شار واکنش‌های شبکه می‌تواند با چندین توزیع شار در شبکه‌ی درشت‌دانه متناظر باشد.

#### ۴-۲-۳ درشت‌دانه‌سازی شبکه

برای هریک از زیرشبکه‌ها، مدهای پایه‌ی آن‌ها را می‌توان با واکنش‌هایی متناظر کرد. ضمناً با توجه به نکات بخش ۴-۲-۲ کل آن زیرشبکه را می‌توان با درشت‌واکنش‌های متناظر جایگزین کرد. در فرایند درشت‌دانه‌سازی شبکه، به جز شبکه‌ی میانی، همه‌ی زیرشبکه‌های موجود با درشت‌واکنش‌های آن‌ها جایگزین می‌شوند و در نتیجه یک شبکه‌ی درشت‌دانه متشکل از درشت‌واکنش‌هایی بین متابولیت‌های میانی به دست می‌آید.

برای مثال با تجزیه‌ی به دست آمده از شکل ۴-۴ حاصل درشت‌دانه‌سازی به صورت شکل ۴-۸ خواهد بود.



شکل ۴-۸ شبکه‌ی درشت‌دانه متناظر با تجزیه‌ی موجود در شکل ۴-۴

هر توزیع شار در این شبکه را می‌توان متشکل از توزیع شار درشت‌واکنش‌ها در هریک زیرشبکه‌ها به همراه توزیع شار برای واکنش‌های میانی دانست. با استفاده از روش گفته شده در بخش ۴-۲-۲، می‌توان برای هریک از زیرشبکه‌ها توزیع شار درشت‌واکنش‌های آن را به توزیع شار برای واکنش‌های آن تبدیل کرد. با توجه به این که بر روی واکنش‌های میانی تبدیلی صورت نگرفته و این واکنش‌ها مستقیماً متناظر با واکنش‌های شبکه‌ی اصلی هستند، شار این واکنش‌ها را مستقیماً برای شبکه‌ی اصلی نیز در نظر گرفت.

#### روش تبدیل توزیع شار در شبکه‌ی درشت‌دانه به توزیع شار متناظر در شبکه‌ی اصلی

برای تبدیل شار شبکه‌ی درشت‌دانه به شار شبکه‌ی اصلی کارهای زیر باید انجام شود:

(۱) برای هریک از زیرشبکه‌ها شار برحسب درشت‌واکنش‌ها را از روش گفته شده در بخش ۴-۲-۲ به شار برحسب واکنش‌های آن زیرشبکه تبدیل می‌کنیم.

(۲) برای واکنش‌های میانی مقدار این واکنش‌ها را مستقیماً برای شبکه‌ی اصلی نیز در نظر می‌گیریم

#### ۴-۲-۴ مدهای پایه‌ی شبکه‌ی درشت‌دانه

با توجه به این که در شبکه‌ی درشت‌دانه هریک از واکنش‌ها در واقع یک مدپایه برای زیرشبکه‌ها بوده‌اند، هر مسیر (یا به عبارتی توزیع شار) در شبکه‌ی درشت‌دانه در واقع یک نحوه‌ی به هم وصل کردن مدهای پایه‌ی زیرشبکه‌ها را با رعایت تعادل استوکیومتری نشان می‌دهد. با توجه به این که هدف ما به دست آوردن مدهای پایه‌ی شبکه‌ی اصلی است، این اتصال باید طوری برقرار گردد که حاصل، نه تنها شرایط استوکیومتری را داشته باشد، بلکه دارای حداقل واکنش‌ها نیز باشد. بنابراین هدف به دست آوردن توزیع‌هایی از شار در شبکه‌ی درشت‌دانه است که پس از تبدیل به توزیع شار در شبکه‌ی اصلی، دارای حداقل واکنش‌های ممکن باشند. هریک از درشت‌واکنش‌ها در زیرشبکه‌ی خود از حداقل واکنش‌ها تشکیل شده‌اند. بنابراین اگر درشت‌مسیرهای ما متشکل از حداقل درشت‌واکنش‌های ممکن باشد انتظار می‌رود که توزیع شار متناظر آن‌ها در شبکه‌ی اصلی نیز دارای حداقل واکنش‌ها باشد. بنابراین رویکرد ما برای به دست آوردن مدهای پایه‌ی شبکه‌ی اصلی به این صورت است که مدهای پایه‌ی شبکه‌ی درشت‌دانه را محاسبه و سپس آن‌ها را برحسب واکنش‌های شبکه‌ی اصلی به دست می‌آوریم.

#### ۴-۲-۵ نسبت مسیرهای به دست آمده با مدهای پایه

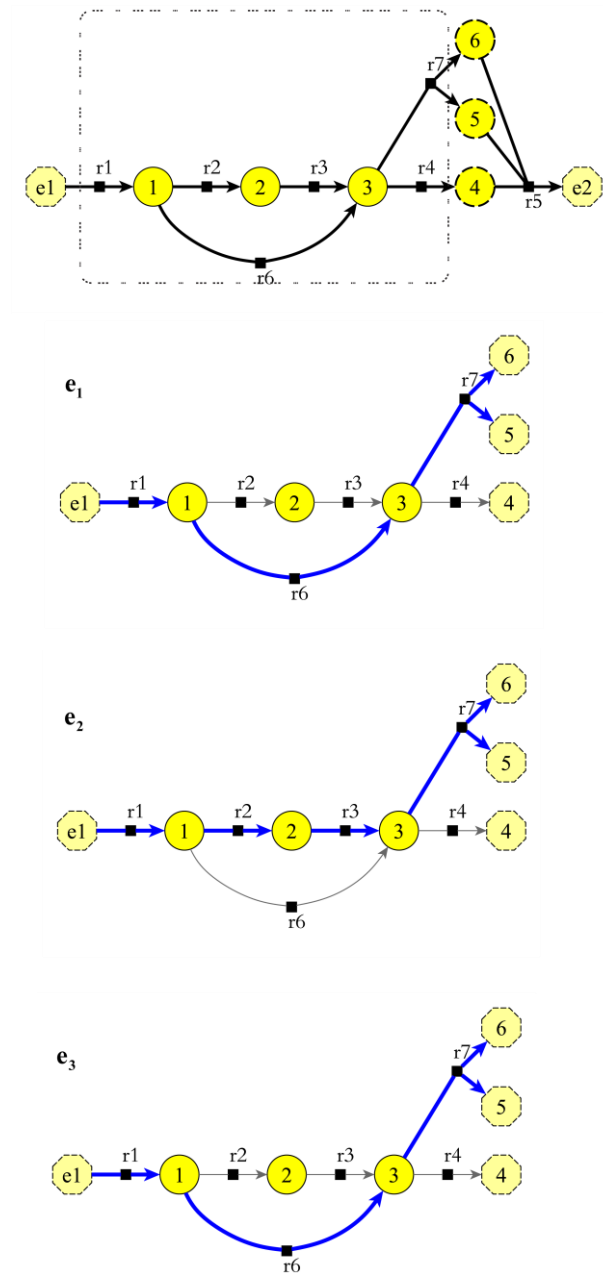
برای این که نشان دهیم بردارهایی که با رویکرد گفته شده در بخش ۴-۲-۴ به دست می‌آیند، شامل همه‌ی مدهای پایه هستند، باید نشان دهیم که هر مد پایه‌ی اصلی، متناظر با حداقل یک مدپایه در شبکه‌ی درشت‌دانه می‌باشد. برای این منظور ابتدا نشان می‌دهیم که اگر یک بردار در شبکه‌ی درشت‌دانه، غیر پایه باشد، در آن صورت معادل آن در شبکه‌ی اصلی نیز غیر پایه خواهد بود.

اگر یک مد شار درشت‌دانه مانند  $e$  غیر پایه باشد، به این معنی است که یک مدشار درشت‌دانه‌ی دیگر مانند  $v$  وجود دارد که  $\text{supp}(v) \subset \text{supp}(e)$ . اگر  $e'$ ، معادل  $e$  در شبکه‌ی اصلی و  $v'$ ، معادل  $v$  در شبکه‌ی اصلی باشد، مجموعه‌ی درشت‌واکنش‌های موجود در  $e$  با اضافه کردن چند واکنش به مجموعه‌ی درشت‌واکنش‌های  $v$  به دست می‌آید. تاوقتی که واکنش‌ها برگشت‌ناپذیر باشند، اضافه کردن درشت‌واکنش‌ها، باعث حذف واکنش‌ها نخواهد شد بنابراین می‌توان گفت  $\text{supp}(v') \subseteq \text{supp}(e')$  و بنابراین  $e'$  نمی‌تواند یک مد پایه باشد.



بنابراین نتیجه می‌گیریم که با به دست آوردن مدهای پایه‌ی شبکه‌ی درشت‌دانه، همه‌ی مدهای پایه‌ی شبکه‌ی اصلی به دست می‌آیند. همچنین با توجه به مثال زیر می‌توان گفت که برخی از بردارهایی که از طریق روش درشت‌دانه سازی به دست می‌آیند، مدپایه نیستند. بنابراین باید از مجموعه‌ی جواب نهایی حذف شوند.

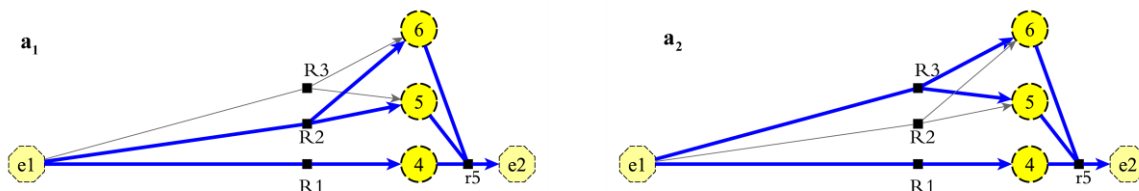
در شبکه‌ی متابولیکی نشان داده شده در ، یک زیرشبکه (شامل متابولیت‌های 1، 2 و 3) مشخص شده و مدهای پایه‌ی آن ( $e_1$ ،  $e_2$  و  $e_3$ ) نشان داده شده‌اند.



شکل ۹-۴ یک زیرشبکه‌ی مشخص شده به همراه مدهای پایهی آن

در شکل ۱۰-۴ دو مد پایهی شبکه‌ی درشت‌دانه ( $\mathbf{a}_1$  و  $\mathbf{a}_2$ ) نشان داده شده‌اند. درشت‌واکنش‌های  $R1$ ،  $R2$  و  $R3$  به ترتیب متناظر با مدهای پایهی  $\mathbf{e}_1$ ،  $\mathbf{e}_2$  و  $\mathbf{e}_3$  در زیرشبکه‌ی مشخص شده در شکل ۹-۴ هستند. مد پایهی  $\mathbf{a}_1$  شامل درشت‌واکنش‌های  $R1$  و  $R2$  است. بنابراین مسیر متناظر در شبکه‌ی اصلی شامل واکنش‌های  $v_1 = \{r1, r2, r3, r4, r5, r7\}$  خواهد بود. مد پایهی  $\mathbf{a}_2$  شامل واکنش‌های  $R1$  و  $R3$  است.

بنابراین مسیر متناظر آن شامل واکنش‌های  $v_2 = \{r1, r2, r3, r4, r5, r6, r7\}$  خواهد بود. با توجه به این که  $v_1 \subset v_2$  بنابراین  $v_2$ ، طبق تعریف نمی‌تواند یک مد پایه برای شبکه‌ی اصلی باشد. زیرا یک بردار شار دیگر ( $v_1$ ) وجود دارد که واکنش‌های آن زیرمجموعه‌ای از واکنش‌های  $v_2$  است.



شکل ۴-۱۰ شبکه‌ی درشت‌دانه برای شبکه‌ی نشان داده شده در شکل ۴-۹

در مورد این مشاهده می‌توان گفت که هنگامی که شبکه‌ی متابولیکی را به صورت درشت‌دانه بررسی می‌کنیم، هریک از درشت‌دانه‌ها به عنوان واکنش‌های مستقل در نظر گرفته می‌شود. درحالی که هریک از واکنش‌ها خود مجموعه‌هایی از واکنش‌ها هستند که ممکن است با یکدیگر اشتراکاتی داشته باشند. این اشتراکات را نمی‌توان در شبکه‌ی درشت‌دانه دید و در نظر گرفت و امکان دارد بعضی از مدهای پایه‌ی شبکه‌ی درشت‌دانه هنگام تبدیل به مسیرهای معادل در شبکه‌ی اصلی، زیرمجموعه‌ی یکدیگر باشند.

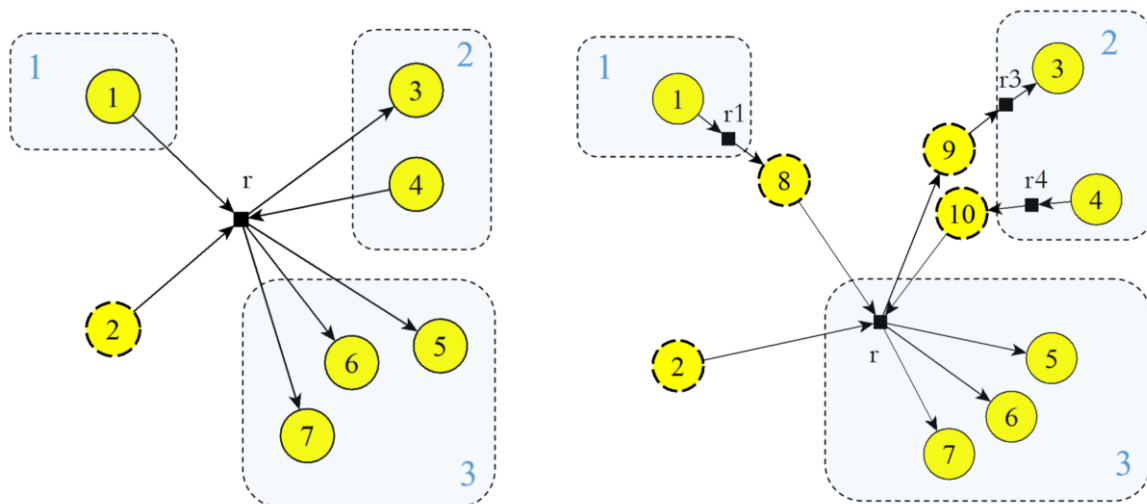
بنابراین این امکان وجود دارد که یک مد پایه برای شبکه‌ی درشت‌دانه، هنگام تبدیل به مسیر معادل در شبکه‌ی اصلی به یک بردار غیر پایه تبدیل شود. این یعنی بردارهایی که از روش درشت‌دانه سازی شبکه به دست می‌آیند شامل بردارهای غیر پایه نیز هستند. البته با توجه به این که این بردارها شامل همه‌ی مدهای پایه هستند، می‌توان با یک فرایند پس پردازش بردارهای غیر پایه را حذف کرده و مجموعه‌ی کامل مدهای پایه را به دست آورد.

## ۴-۳ به وجود آوردن شرط جدایی زیرشبکه‌ها با افزودن گره‌های ساختگی

روش معرفی شده در بخش ۴-۱-۲ برای عمل درشت‌دانه‌سازی، به یک افراز از متابولیت‌های شبکه نیاز دارد که در آن بین متابولیت‌های زیرشبکه‌ها هیچ واکنشی وجود نداشته باشد. اما مشکلی که وجود دارد این است که برای یک شبکه‌ی دلخواه داده شده، ممکن است یک افراز با این شرایط وجود نداشته باشد.

حتی در صورتی که این کار ممکن باشد، پیدا کردن چنین افزایی ممکن است کار ساده‌ای نباشد. ضمن این که این شرایط باعث می‌شود که انعطاف زیادی در انتخاب زیرشبکه‌ها وجود نداشته باشد.

برای مقابله با این مشکلات، ما از رویکرد اضافه کردن گره‌های ساختگی استفاده می‌کنیم. در این روش ابتدا متابولیت‌ها به طور دلخواهی افزای می‌گردند. البته این افزای به گونه‌ای صورت می‌گیرد که متابولیت‌های گروه‌های مختلف کم‌ترین ارتباطات ممکن را با یکدیگر داشته باشند. سپس به منظور جدا کردن زیرشبکه‌های متصل، متابولیت‌های ساختگی بین زیر شبکه‌ها قرار می‌گیرند و به این ترتیب اتصال آن‌ها به واسطه‌ی متابولیت ساختگی میانجی‌گری می‌شود. این کار به این منظور انجام می‌گیرد که پس از اضافه کردن تعداد مناسبی متابولیت‌های ساختگی بتوان اطمینان حاصل کرد که هیچ دو متابولیتی از زیرشبکه‌های مختلف در یک واکنش مشترک نیستند. در نتیجه‌ی این تغییر یک شبکه‌ی افزوده به دست می‌آید که دارای یک تجزیه‌ی واجد شرایط می‌باشد. البته تناظر یک‌به‌یکی بین فضای شار این شبکه‌ی افزوده و شبکه‌ی اصلی برقرار است. بنابراین می‌توان فرایند محاسبه‌ی مدهای پایه‌ی زیر شبکه‌ها و سپس ساختن مدهای پایه‌ی کلی را در این شبکه دنبال کرد. با توجه به وجود تناظر بین شبکه‌ی افزوده و شبکه‌ی اصلی، می‌توان مطمئن بود که پس از به دست آوردن مدهای پایه‌ی برای شبکه‌ی افزوده می‌توان به راحتی آن‌ها را به مدهای پایه‌ی شبکه‌ی اصلی تبدیل کرد.



شکل ۴-۱۱ نحوه‌ی افزودن متابولیت‌های ساختگی برای هر یک از واکنش‌های بین زیرشبکه‌ای

هر واکنش بین زیرشبکه‌ای باید مورد بررسی قرار گیرد و تعدادی متابولیت ساختگی اضافه شوند. نحوه‌ی اضافه کردن متابولیت‌های ساختگی به این صورت است که ابتدا باید واکنش مورد نظر به یکی از

زیرشبکه‌هایی که با آن ارتباط دارد نسبت داده شود. سپس به ازای هریک از متابولیت‌های زیرشبکه‌های دیگر، یک متابولیت ساختگی به این واکنش متصل می‌شوند.

برای مثال واکنش  $r$  از زیرشبکه‌ی ۱ به متابولیت ۱، از زیرشبکه‌ی ۲ به متابولیت‌های ۳ و ۴ و از زیرشبکه‌ی ۳ به متابولیت‌های ۵ و ۶ و ۷ متصل است. همچنین این واکنش به متابولیت میانی ۲ نیز متصل است. عملیات اضافه کردن متابولیت‌های ساختگی به این صورت است که ابتدا واکنش  $r$  را به یکی از زیرشبکه‌ها مثلاً ۳ نسبت می‌دهیم. سپس به ازای هریک از متابولیت‌های زیرشبکه‌های دیگر باید یک متابولیت ساختگی اضافه شود. در این مثال برای متابولیت‌های ۱ و ۳ و ۴ به ترتیب متابولیت‌های ۸، ۹ و ۱۰ اضافه شده است. ضریب استوکیومتری واکنش  $r$  برای هریک از متابولیت‌ها عیناً به متابولیت ساختگی مربوطه منتقل می‌شود. برای مثال اگر ضریب استوکیومتری  $r$  برای متابولیت ۴ در شبکه‌ی اصلی برابر ۲- باشد، در شبکه‌ی جدید باید ضریب  $r$  برای متابولیت ۱۰ نیز برابر ۲- باشد. ضمن این که برای اتصال هریک از متابولیت‌های ساختگی به متابولیت اصلی یک واکنش باید اضافه شود. این واکنش‌های ساختگی دارای جهتی منطبق بر جهت واکنش  $r$  نسبت به متابولیت مربوطه است. مثلاً به دلیل این که واکنش  $r$  متابولیت ۴ را مصرف می‌کند، واکنش ساختگی اضافه شده نیز باید متابولیت ۴ را مصرف کند. همچنین برگشت پذیری واکنش  $r$  به همه‌ی واکنش‌های ساختگی حاصل منتقل می‌شود یعنی اگر واکنش  $r$  برگشت پذیر باشد، همه‌ی واکنش‌های  $r1$ ،  $r3$  و  $r4$  نیز باید برگشت‌پذیر تعریف شوند.

اضافه کردن متابولیت‌ها و واکنش‌های ساختگی هیچ تاثیری روی شارهای مجاز شبکه ندارد و هر شار به دست آمده برای شبکه‌ی افزوده را می‌توان تنها با حذف مقادیر مربوط به واکنش‌های ساختگی به یک شار برای شبکه‌ی اصلی تبدیل کرد.

## ۴-۴ مسیرهای موازی

نتایج فرایند درشت‌دانه سازی شبکه نشان می‌دهد که پس از این عملیات، تعداد قابل توجهی واکنش‌های موازی به وجود می‌آید. منظور از واکنش‌های موازی واکنش‌هایی است که دقیقاً اثر یکسانی دارند و ستون متناظر با آن‌ها در ماتریس استوکیومتری دقیقاً یکسان است. برای مثال در شبکه‌ی شکل

۴-۹ دو مدپایه  $e_2$  و  $e_3$  دو مسیر موازی می‌باشند. زیر از مسیرهای مختلفی گذشته اما اثرات نهایی آن‌ها یکسان است. این دو مسیر موازی هنگام درشت‌دانه‌سازی به درشت‌واکنش‌های موازی  $R2$  و  $R3$  تبدیل می‌شوند که این درشت‌واکنش‌ها به راحتی می‌توانند به صورت یک واکنش فشرده گردند. در هنگام بازگرداندن مدهای پایه‌ی شبکه‌ی درشت‌دانه به شبکه‌ی اصلی، درشت‌واکنش‌های موازی که به صورت یک درشت‌واکنش فشرده شده‌اند، مجدداً گسترش می‌یابند. مسیرهایی که از درشت‌واکنش فشرده‌شده عبور می‌نمایند، بر حسب این که از  $R2$  یا  $R3$  عبور کنند، به دو مسیر متناظر می‌گردند.

مسیرهای موازی در شبکه‌های متابولیکی احتمالاً از افزونگی ناشی می‌شوند. یعنی برای یک عملکرد متابولیکی مشخص ممکن است مسیرهای متعددی داشته باشد. این مسیرها به قوام شبکه‌های متابولیکی کمک می‌کنند یعنی در صورت مسدود شدن یک مسیر متابولیکی، شبکه‌ی متابولیکی می‌تواند یک فرایند خاص را از یک مسیر دیگر صورت دهد. صرف نظر از دلیل یا مفهوم زیستی مسیرهای موازی، با توجه به و در فرایند درشت‌دانه‌سازی شبکه مسیرهای موازی می‌توانند به (درشت‌واکنش‌های موازی تبدیل شوند و از این خاصیت برای فشرده‌سازی استفاده گردد. در ادامه جزئیات این عملیات فشرده‌سازی و گسترش (درشت‌واکنش‌های موازی توضیح داده شده است.

### فشرده‌سازی واکنش‌های موازی

اگر در شبکه‌ی متابولیکی  $N$  دو واکنش  $r1$  و  $r2$  موازی باشند، در این صورت برای محاسبه‌ی مدهای پایه‌ی شبکه، می‌توان دو واکنش  $r1$  و  $r2$  را باهم در یک واکنش  $r$  یکی کرد تا شبکه‌ی  $N'$  حاصل شود. پس از به دست آوردن مدهای پایه‌ی شبکه‌ی  $N'$ ، به صورت زیر می‌توان مدهای پایه‌ی شبکه‌ی  $N$  را با گسترش نسبت به واکنش  $r$  محاسبه کرد.

۱. در مورد مدهایی که شار واکنش  $r$  در آن‌ها صفر است، می‌توان این مدها را مستقیماً به مدهای شبکه‌ی  $N$  تبدیل کرد. به این صورت که شارهمه‌ی واکنش‌های دیگر مشابه و شار  $r1$  و  $r2$  نیز صفر در نظر گرفته می‌شود.

۲. در مورد مدهایی که شار واکنش  $r$  در آن‌ها غیر صفر است، در شبکه‌ی  $N$  متناظر با آن‌ها دو مدپایه وجود خواهد داشت. مدپایه‌ای که در آن شار واکنش  $r1$  برابر شار  $r$  است و شار  $r2$  برابر صفر است. و مدپایه‌ای که در آن شار واکنش  $r1$  برابر صفر و شار واکنش  $r2$  برابر شار واکنش  $r$  است.

همچنین فشردن سازی (درشت) واکنش‌های موازی به سادگی می‌تواند برای هر تعداد واکنش موازی تعمیم یابد.

در واقع تبدیل مسیرهای موازی به واکنش‌های موازی بخشی از انگیزه‌ی تجزیه‌ی شبکه در این کار را تشکیل می‌دهد. این مشاهده که شبکه‌های متابولیکی از ماژول‌هایی تشکیل شده‌اند که دارای ارتباطات داخلی زیاد و ارتباطات خارجی کم هستند این احتمال را به وجود می‌آورند که تعداد زیادی از مسیرهایی که از این زیرشبکه‌ها می‌گذرند با هم معادل هستند. اگر شبکه‌ی متابولیکی به صورت یکپارچه بررسی شود، امکان فشردن سازی مسیرهای موازی وجود ندارد. اما فرآیند تبدیل مسیرهای زیرشبکه‌ها به واکنش‌های معادل آن‌ها این امکان را فراهم می‌کند که با یکی کردن واکنش‌های معادل، فشردن سازی صورت گیرد.

## ۴-۵ روند کلی روش تجزیه به ازای یک افراز دلخواه از شبکه و

### محاسبه‌ی مدهای پایه

در بخش‌های گذشته هریک از مراحل روش محاسبه‌ی مدهای پایه از طریق تجزیه‌ی شبکه به تفصیل شرح داده شد. در این بخش روند کلی به صورت ترتیبی ارائه می‌شود.

۱. فرایند تجزیه‌ی شبکه‌ی متابولیکی از افراز مجموعه‌ی متابولیت‌ها به چند زیرمجموعه آغاز می‌شود. که یکی از این زیرمجموعه‌ها مجموعه‌ی متابولیت‌های میانی است. در این افراز باید حداکثر سعی انجام گیرد تا متابولیت‌های زیرشبکه‌های مختلف کمترین ارتباط را با یکدیگر داشته باشند و همه‌ی ارتباطات از طریق متابولیت‌های میانی صورت گیرد.

۲. اگر بین متابولیت‌های موجود در گروه‌های مختلف واکنشی وجود داشته باشد، به ازای هر واکنش بین زیرشبکه‌ای لازم است یک یا چند متابولیت ساختگی به مجموعه‌ی متابولیت‌های میانی اضافه شود تا ارتباط آن‌ها با واسطه‌ی این متابولیت‌های ساختگی برقرار گردد. نتیجه‌ی این فرایند یک شبکه‌ی افزوده است که افراز واجد شرایط بلافاصله برای آن به دست می‌آید.

۳. با استفاده از تجزیه‌ی واجد شرایط به دست آمده از مرحله‌ی ۱ یا ۲، عملیات درشت‌دانه‌سازی شبکه می‌تواند انجام گیرد. نتیجه‌ی این مرحله یک شبکه‌ی درشت‌دانه است که شامل متابولیت‌های میانی، متابولیت‌های خارجی و واکنش‌هایی بین آن‌ها است. واکنش‌های شبکه‌ی درشت‌دانه در اصل یا واکنش‌هایی بین متابولیت‌های میانی بوده‌اند یا درشت‌واکنش‌هایی هستند که از مدهای پایه‌ی زیرشبکه‌ها به دست آمده‌اند.

۴. در شبکه‌ی درشت‌دانه، واکنش‌های موازی به صورت یک واکنش فشرده می‌شوند. اما باید اطلاعات مربوط به هریک از این واکنش‌های موازی نگه داشته شود. یعنی باید مشخص شود که واکنش به دست آمده حاصل فشرده‌کردن چه مدهای پایه‌ای است.

۵. در مرحله‌ی بعد مدهای پایه‌ی شبکه‌ی درشت‌دانه به دست می‌آیند. این مدهای پایه‌ی درشت‌دانه، بردارهای شار برحسب درشت‌واکنش‌ها و واکنش‌های میانی هستند. این بردارهای شار باید به بردارهای معادل برحسب واکنش‌های شبکه‌ی اصلی تبدیل شوند. این عمل برای هریک از زیرشبکه‌ها با استفاده از جمع وزن دار مدهای شار آن زیرشبکه به دست می‌آید. ضریب هریک از این مدهای شار برابر شار درشت‌واکنش معادل در بردار شار شبکه‌ی درشت‌دانه است. برای واکنش‌های میانی شار شار آن‌ها مستقیماً برای شبکه‌ی اصلی نیز در نظر گرفته می‌شود.

۶. در میان بردارهای به دست آمده ممکن است بردارهای تکراری و بردارهای غیر پایه وجود داشته باشد. بنابراین باید این بردارهای اضافه با یک فرایند پس‌پردازش حذف شوند. فرایند پس‌پردازش به این صورت است که هر برداری که یک زیر مجموعه از آن (از لحاظ واکنش‌های موجود) در میان مجموعه‌ی بردارهای به دست آمده وجود داشته باشد، باید حذف گردد.

۷. در صورت اضافه شدن متابولیت‌ها و واکنش‌های ساختگی (از مرحله‌ی ۲)، با حذف مولفه‌های مربوط به واکنش‌های ساختگی از بردارهای به دست آمده، این بردارها به بردارهایی برای شبکه‌ی اصلی تبدیل می‌شوند.

بردارهایی که از مرحله‌ی ۷ به دست می‌آیند، باید بر مجموعه‌ی مدهای پایه‌ی شبکه منطبق باشند.



## ۴-۶ نحوه‌ی انتخاب زیرشبکه‌ها

با فراهم شدن امکان اضافه کردن متابولیت‌ها و واکنش‌های ساختگی، روش معرفی شده می‌تواند برای هر شبکه‌ی دلخواهی به کار رود. برای یک شبکه‌ی دلخواه داده شده، با مشخص کردن یک افراز دلخواه از گره‌ها و انتخاب یکی از زیرمجموعه‌های حاصل به عنوان مجموعه‌ی گره‌های میانی می‌توان عملیات محاسبه‌ی مدهای پایه از طریق تجزیه‌ی شبکه را برای هر شبکه‌ای انجام داد. البته در این فرآیند ممکن است تعدادی گره ساختگی به شبکه اضافه شوند. در واقع به ازای هر واکنش بین زیر شبکه‌ای یک یا چند گره (و واکنش) اضافی به شبکه اضافه خواهد شد. طبیعی است که هرچه واکنش‌های بین زیرشبکه‌ای بیشتر باشند، تعداد بیشتری گره اضافی نیز مورد نیاز خواهد بود. اضافه کردن گره‌های ساختگی این امکان را به ما می‌دهد تا به جای محدود کردن خود به نوع خاصی از افراز، بتوان افراز را با یک هدف خاص انجام داده و سپس بتوان آن را به یک تجزیه‌ی واجد شرایط تبدیل کرد. در این صورت می‌توان گفت که تعداد متابولیت‌های ساختگی اضافه شده به نحوی کیفیت افراز ما را برای فرایند تجزیه و درشت‌دانه سازی مشخص می‌سازد. زیرا این متابولیت‌ها و واکنش‌های ساختگی با افزایش اندازه‌ی شبکه‌ی درشت‌دانه باعث ایجاد سربار در روش می‌گردند.

بنابراین باید سعی کرد افراز متابولیت‌ها طوری انجام شود که مجموعه‌های حاصل تا حد امکان ویژگی‌های زیر را داشته باشند:

۱. زیرمجموعه‌های غیر از مجموعه‌ی متابولیت‌های میانی، تا حد امکان جدا از هم باشند. یعنی کمترین تعداد واکنش بین آن‌ها وجود داشته باشد.

۲. مجموعه‌ی متابولیت‌های میانی طوری انتخاب شوند که بتوانند به خوبی نقش واسط بین زیرشبکه‌های دیگر را ایفا کنند. ضمن این که تعداد این متابولیت‌ها حداقل ممکن باشد.

رعایت این ویژگی‌ها باعث می‌شود که تجزیه‌ای با کیفیت‌تر یعنی تجزیه‌ای با کمترین تعداد متابولیت ساختگی به دست آید.

برای شبکه‌های کوچک ممکن است افراز متابولیت‌ها به روش دستی امکان‌پذیر باشد. اما برای شبکه‌های بزرگ باید یک فرایند خودکار برای عمل افراز در نظر گرفته شود. میزان ارتباطات بین زیرشبکه‌ها را

می‌توان با معیار اندازه‌ی برش<sup>۱</sup> بیان کرد. این معیار برای ابرگراف، که به خوبی ساختار ارتباطات را در شبکه‌های متابولیکی نشان می‌دهد نیز تعریف شده است. مساله‌ی یافتن افرازی با حداقل اندازه‌ی برش برای یک ابرگراف یک مساله‌ی سخت (از دسته‌ی NP سخت) است. بنابراین الگوریتم کارامدی که جواب دقیق این مساله را در زمان معقول بیابد وجود ندارد. اما الگوریتم‌های ابتکاری زیادی به خصوص در حوزه‌ی طراحی مدارهای پرتراکم برای این مساله به وجود آمده که قادرند با سرعت بسیار بالا جواب‌های با کیفیتی برای ابرگراف‌های بزرگ به دست آورند. بنابراین این الگوریتم‌ها گزینه‌ی برای افراز متابولیت‌های غیرمیانی هستند. البته قبل از آن باید متابولیت‌های میانی تعیین و از شبکه حذف گردند. تعیین و حذف اولیه‌ی متابولیت‌های میانی با این توجیه انجام می‌گیرد که طبق ساختار ماژولار سلسله‌مراتبی که برای شبکه‌های متابولیکی پیشنهاد شده، در این شبکه‌ها معمولاً بعضی متابولیت‌ها نقش مرکزی دارند و به تعداد زیادی از متابولیت‌های دیگر وصل هستند. بنابراین بهتر است که به یک زیرشبکه‌ی خاص تعلق نگیرند بلکه به عنوان متابولیت میانی در نظر گرفته شوند و نقش واسط بین زیر شبکه‌ها را داشته باشند. متابولیت‌های دیگر که اتصالات محلی‌تر دارند و به تعداد زیادی از متابولیت‌های دیگر متصل نیستند، بهتر است به همراه متابولیت‌های مجاور خود در داخل یکی از زیر شبکه‌ها قرار گیرند.

بر همین مبنا، در چارچوب معرفی شده توسط شوستر از این ویژگی برای افراز شبکه استفاده شده است. در واقع در این چارچوب، با در نظر گرفتن یک آستانه‌ی ۴ برای درجه‌ی گره‌ها سعی در یافتن متابولیت‌های مرکزی در ساختار متابولیسم می‌شود [۵۰]. در روش ارائه شده در این تحقیق، در راستای چارچوب شوستر، ابتدا متابولیت‌های میانی بر مبنای آستانه‌ی ۴ برای درجه‌ی گره یافته می‌شوند. سپس این متابولیت‌ها از شبکه حذف و ابرگراف باقی مانده با هدف کمینه‌کردن اندازه‌ی برش افراز می‌گردد. برای افراز مجموعه‌ی متابولیت‌ها می‌توان از الگوریتم hMETIS استفاده نمود. این الگوریتم که از بهترین الگوریتم‌ها در زمینه‌ی افراز ابرگراف است، امکان افراز به تعداد دلخواهی زیرشبکه را نیز دارد [۵۸].

---

<sup>۱</sup> Cut Size

بنابراین می‌توان یک رویه‌ی ثابت برای تجزیه‌ی مناسب شبکه‌های متابولیکی به این صورت در نظر گرفت:

۱. ابتدا یک زیرمجموعه از متابولیت‌ها به عنوان مجموعه‌ی میانی انتخاب می‌شوند. این انتخاب می‌تواند همانند روش‌های قبلی با در نظر گرفتن یک حد آستانه‌ی ۴ برای درجه‌ی گره‌ها یا با هر معیار دیگری انتخاب گردد.

۲. پس از تعیین متابولیت‌های میانی، این متابولیت‌ها حذف می‌شوند. (بدون حذف واکنش‌های متصل به آن‌ها). سپس شبکه‌ی باقی مانده با هدف کاهش ارتباطات بین زیر شبکه‌ها، به تعدادی زیرشبکه افراز می‌شود.

با توجه به این که ممکن است از قبل تعداد مناسب برای زیرشبکه‌ها مشخص نباشد، می‌توان تعداد مختلف پارتیشن‌ها را آزمایش کرد و بهترین نتیجه را که معادل با کمترین تعداد متابولیت‌های ساختگی است انتخاب نمود.

بنابراین با دریافت یک شبکه‌ی متابولیکی دلخواه، با استفاده از مراحل ۱ و ۲ می‌توان یک افراز برای متابولیت‌های شبکه به دست آورد که در آن متابولیت‌های با درجه‌ی بالا به عنوان متابولیت‌میانی مشخص شده‌اند و متابولیت‌های دیگر با هدف کمینه‌کردن اتصالات بین آن‌ها افراز گردیده‌اند. پس از اضافه کردن متابولیت‌های میانی در محل اتصال زیرشبکه‌ها، می‌توان یک تجزیه‌ی واجد شرایط برای شبکه به دست آورد و مراحل درشت‌دانه‌سازی و محاسبه‌ی مدهای پایه را به دنبال آن انجام داد.

۵

فصل پنجم

نتایج و تحلیل آن‌ها

## نتایج و تحلیل آن‌ها

### ۵-۱ شرایط آزمایش

الگوریتم ارایه شده با استفاده از نرم‌افزار MATLAB R2015b پیاده‌سازی شده و اجرای الگوریتم بر روی شبکه‌های نمونه‌ی واقعی نیز با استفاده از همین نرم‌افزار صورت گرفته است. این پیاده‌سازی، شبکه‌ی متابولیکی را به صورت ماتریس استوکیومتری به همراه مجموعه‌ی واکنش‌های برگشت‌ناپذیر دریافت می‌نماید. همچنین یک افراز روی مجموعه‌ی متابولیت‌ها به الگوریتم داده می‌شود. سپس عملیات تشکیل زیرشبکه‌ها و درشت‌دانه‌سازی آن‌ها انجام می‌گیرد. در همه‌ی موارد آزمایش، زمان فرایند تبدیل شبکه به نسبت زمان محاسبه‌ی مدهای پایه قابل چشم‌پوشی بوده و بنابراین نیاز به یک پیاده‌سازی با استفاده از زبان‌های سطح پایین‌تر احساس نمی‌شود.

برای این که فرایند افراز متابولیت‌ها به صورت خودکار انجام شود، از الگوریتم hMETIS استفاده شده است [۵۸]. این الگوریتم به صورت بسته‌ای از چند برنامه‌ی تحت ویندوز ارائه شده است که از طریق خط فرمان اجرا می‌گردند. برنامه‌ی khmetis برای افراز گره‌های یک ابرگراف به  $k$  زیرابگراف استفاده می‌گردد. این برنامه پارامترهایی جهت تعیین نحوه‌ی عمل افراز، به همراه یک فایل نوشتاری که شامل توصیف ابرگراف مورد نظر است دریافت می‌نماید. نتیجه‌ی فرایند افراز از طریق یک فایل نوشتاری از این برنامه دریافت می‌گردد.

آزمایش‌ها بر روی یک کامپیوتر با مشخصات زیر اجرا شده است.

CPU: Intel Core-i5 6300U 2.4GHz

Memory: 4GB

OS: Windows 10 Pro

محاسبه‌ی مدهای پایه برای شبکه‌ی اصلی و زیرشبکه‌ها با استفاده از برنامه‌ی efmtool انجام گرفته که این برنامه به زبان جاوا نوشته شده و از داخل MATLAB فراخوانی می‌شود [۲۵].

## ۵-۲ نتایج اجرا

برای بررسی عملکرد الگوریتم، برنامه‌ی پیاده‌سازی شده، بر روی ۴ شبکه‌ی متابولیکی واقعی در اندازه‌های مختلف اعمال گردیده است. در هریک از آزمایش‌های ۱ تا ۴، نتایج تجزیه و درشت‌دانه‌سازی یکی از این شبکه‌ها، به ازای تعداد مختلف زیرشبکه‌ها، در جدول مربوطه آورده شده است. برای آزمایش‌های ۱ و ۲ که محاسبه‌ی مدهای پایه ممکن بوده، زمان عملیات و تعداد بردارهای حاصل شده و برای آزمایش‌های ۳ و ۴ که محاسبه‌ی آن‌ها امکان‌پذیر نبوده، تنها اندازه‌ی شبکه‌ی درشت‌دانه و اندازه‌ی ماتریس کرنل برای مقایسه با شبکه‌ی اصلی گزارش شده است. اندازه‌ی ماتریس کرنل به عنوان شاخص اصلی برای پیچیدگی محاسبه‌ی مدهای پایه در مراجع مختلفی ذکر شده است [۲۴، ۲۵]. با توجه به این که تعداد پیمایش‌های حلقه‌ی اصلی در الگوریتم پوچ‌فضا، ارتباط مستقیمی با اندازه‌ی ماتریس کرنل دارد، این معیار می‌تواند برای ارزیابی زمانی که این الگوریتم طول خواهد کشید به کار رود.

در هر یک از جدول‌ها سطر مربوط به بهترین حالت تجزیه با زمینه‌ی خاکستری مشخص شده است. برای آزمایش ۱ و ۲ معیار انتخاب بهترین تجزیه زمان اجرا بوده و برای آزمایش ۳ و ۴ معیار انتخاب بهترین تجزیه، اندازه‌ی ماتریس کرنل در نظر گرفته شده است.

در حال حاضر روش‌هایی موسوم به فشرده‌سازی شبکه<sup>۱</sup> یا کاهش شبکه<sup>۲</sup> برای شبکه‌های متابولیکی وجود دارد که این روش‌ها عمدتاً از ویژگی‌های روابط پایستگی<sup>۳</sup> و زیرمجموعه‌ی آنزیم‌ها<sup>۴</sup> استفاده می‌نمایند [۱۳، ۱۴، ۲۴]. این روش‌های به صورت یک برنامه‌ی MATLAB به همراه نرم‌افزار Metatool در دسترس می‌باشد. در این جا این برنامه بر روی هریک از شبکه‌های مورد آزمایش اعمال گردیده و در مورد هر آزمایش نتایج آن با نتایج روش درشت‌دانه‌سازی شبکه مقایسه شده است.

<sup>۱</sup> Network Compression

<sup>۲</sup> Network Reduction

<sup>۳</sup> Conservation Relation

<sup>۴</sup> Enzyme Subset

آزمایش ۱ (Glycolysis and Pentose Phosphate Pathway in T.brucei [۵۹]): در این مرحله ابتدا نتایج اجرای الگوریتم بر روی یک شبکه‌ی متابولیکی کوچک با ۳۹ متابولیت و ۴۰ واکنش گزارش شده است (جدول ۵-۱).

جدول ۵-۱ نتایج آزمایش ۱

| تعداد زیر شبکه‌ها | تعداد متابولیت‌های ساختگی | تعداد متابولیت‌ها پس از درشت‌دانه سازی | تعداد واکنش‌ها پس از درشت‌دانه سازی | تعداد واکنش‌ها پس از یکی کردن درشت‌واکنش‌های موازی | تعداد ستون‌های کرنل پس از یکی کردن درشت‌واکنش‌های موازی | تعداد مدهای پایه‌ی شبکه‌ی درشت‌دانه پس از یکی کردن واکنش‌های موازی | زمان محاسبه efmtool (میلی ثانیه) | تعداد پیمایش‌های حلقه‌ی اصلی در efmtool |
|-------------------|---------------------------|--|-------------------------------------|--|---|--|----------------------------------|---|
| ۲                 | ۱                         | ۵                                      | ۲۰                                  | ۱۴   | ۱۱  | ۴۴   | ۴۰-۶۰                            | ۴                                       |
| ۳                 | ۲                         | ۶                                      | ۲۷                                  | ۱۸   | ۱۴  | ۱۲۶  | ۴۰-۶۰                            | ۴                                       |
| ۴                 | ۴                         | ۸                                      | ۲۷                                  | ۱۸   | ۱۲  | ۱۳۱  | ۶۰-۸۰                            | ۵                                       |
| ۵                 | ۴                         | ۸                                      | ۳۲                                  | ۲۴   | ۱۸  | ۲۳۱۲   | ۱۲۰-۱۶۰                          | ۵                                       |
| ۶                 | ۵                         | ۹                                      | ۲۵                                  | ۲۲   | ۱۵  | ۶۵۸  | ۱۳۰-۱۴۰                          | ۷                                       |
| ۷                 | ۱۲                        | ۱۶                                     | ۳۷                                  | ۳۴   | ۲۰  | ۲۳۲۸۴  | ۸۰۰-۹۰۰                          | ۱۰                                      |
| ۸                 | ۱۲                        | ۱۶                                     | ۳۴                                  | ۳۲   | ۲۰  | ۵۰۷۹   | ۲۵۰-۳۰۰                          | ۱۰                                      |
|                   |                           | متابولیت‌های شبکه‌ی اصلی (۳۹)          | واکنش‌های شبکه‌ی اصلی (۴۰)          | واکنش‌های شبکه‌ی اصلی (۴۰)                         | واکنش‌های شبکه‌ی اصلی (۴۰)                              | ستون‌های کرنل در شبکه‌ی اصلی (۸)                                   | مدهای شبکه‌ی اصلی (۱۲۰) ثانیه    | زمان شبکه‌ی اصلی ۱۱۸-۱۳۳ میلی ثانیه     |

پس از اعمال روش‌های فشرده‌سازی موجود، این شبکه به شبکه‌ای با ۱۴ متابولیت و ۱۷ واکنش و ماتریس کرنل با ۸ ستون تبدیل شده است. در مقابل با استفاده از روش درشت‌دانه‌سازی در بهترین حالت شبکه‌ای با ۵ متابولیت و ۱۴ واکنش و ماتریس کرنل با ۱۱ ستون حاصل شده است.

عملیات به ازای تعداد مختلفی از زیر شبکه‌ها (۲ تا ۸ زیر شبکه) انجام گرفته و نتایج مراحل مختلف در جدول آمده است. در مرحله‌ی محاسبه‌ی مدهای پایه توسط efmtool، به دلیل این که زمان اجرا ممکن است در دفعات مختلف کمی متفاوت باشد، برای هر مورد، آزمایش ۵ بار تکرار شده و بازه‌ی تغییرات

گزارش شده است. به علاوه، در همه‌ی موارد انطباق جواب‌های به دست آمده (پس از گسترش مربوط به واکنش‌های موازی و حذف بردارهای غیر پایه) با جواب‌های روش عادی بررسی و تایید شده است.



## آزمایش ۲ (E.coli CORE [۶۰])

در جدول ۲-۵ نتایج آزمایش بر روی یک مدل متابولیسم با اندازه‌ی متوسط (۷۲ متابولیت و ۹۴ واکنش) نشان داده شده است.

جدول ۲-۵ نتایج آزمایش ۲

| تعداد زیر شبکه‌ها | تعداد متابولیت‌های ساختگی | تعداد متابولیت‌های سازش از درشت‌دانه سازش | تعداد واکنش‌ها پس از درشت‌دانه سازش | تعداد واکنش‌ها پس از یکی کردن درشت‌واکنش‌های موازی | تعداد واکنش‌های کرنل پس از یکی کردن درشت‌واکنش‌های موازی | تعداد مدهای پایه شبکه‌ی درشت‌دانه پس از یکی کردن واکنش‌های موازی | زمان محاسبه efmtool روی شبکه‌ی تبدیل یافته (میلی ثانیه) | تعداد پیمایش‌های حلقه‌ی اصلی در efmtool |
|-------------------|---------------------------|---|-------------------------------------|--|--|--|---|---|
| ۲                 | ۰                         | ۲۱  | ۴۸                                  | ۴۴   | ۲۷   | ۶۳۴۱۲  | ۵۷۴۲  | ۱۶                                      |
| ۴                 | ۰                         | ۲۱  | ۴۸                                  | ۴۴   | ۲۷   | ۶۳۴۱۲  | ۶۲۷۹  | ۱۶                                      |
| ۴                 | ۰                         | ۲۱  | ۴۸                                  | ۴۴   | ۲۷   | ۶۳۴۱۲  | ۴۷۰۹  | ۱۶                                      |
| ۵                 | ۰                         | ۲۱  | ۴۸                                  | ۴۴   | ۲۷   | ۶۳۴۱۲  | ۴۳۴۵  | ۱۶                                      |
| ۶                 | ۰                         | ۲۱  | ۴۸                                  | ۴۴   | ۲۷   | ۶۳۴۱۲  | ۴۴۳۱  | ۱۶                                      |
| ۷                 | ۱                         | ۲۲  | ۴۹                                  | ۴۵   | ۲۷   | ۸۸۲۳۰  | ۷۹۵۲  | ۱۷                                      |
| ۸                 | ۱                         | ۲۲  | ۴۹                                  | ۴۵   | ۲۷   | ۸۸۲۳۰  | ۶۹۱۴  | ۱۷                                      |
| ۹                 | ۱                         | ۲۲  | ۴۹                                  | ۴۵   | ۲۷   | ۶۳۴۱۲  | ۳۰۵۳  | ۱۶                                      |
| ۱۰                | ۱                         | ۲۲  | ۴۹                                  | ۴۵   | ۲۷   | ۶۳۴۱۲  | ۴۲۹۹  | ۱۶                                      |
| ۱۱                | ۱                         | ۲۲  | ۴۹                                  | ۴۵   | ۲۷   | ۶۳۴۱۲  | ۴۵۳۰  | ۱۶                                      |
| ۱۲                | ۲                         | ۲۳  | ۵۰                                  | ۴۶   | ۲۷   | ۶۳۴۱۲  | ۴۳۳۳  | ۱۶                                      |
| ۱۳                | ۴                         | ۲۵  | ۵۳                                  | ۴۹   | ۲۸   | ۱۴۵۰۳۴   | ۹۱۱۱  | ۱۹                                      |
| ۱۴                | ۳                         | ۲۴  | ۵۰                                  | ۴۶   | ۲۶   | ۵۵۳۵۵  | ۳۹۹۷  | ۱۷                                      |
| ۱۵                | ۵                         | ۲۶  | ۵۳                                  | ۴۹   | ۲۷   | ۱۲۵۴۸۱   | ۸۴۳۳  | ۲۰                                      |
| ۱۶                | ۵                         | ۲۶  | ۵۳                                  | ۴۹   | ۲۷   | ۱۲۵۴۸۱   | ۱۱۵۳۱   | ۲۰                                      |
| ۱۷                | ۵                         | ۲۶  | ۵۳                                  | ۴۹   | ۲۷   | ۱۲۵۴۸۱   | ۷۸۸۶  | ۲۰                                      |
| ۱۸                | ۵                         | ۲۶  | ۵۴                                  | ۵۰   | ۲۸   | ۱۹۵۶۰۷   | ۱۲۱۴۵   | ۲۰                                      |
| ۱۹                | ۷                         | ۲۸  | ۵۶                                  | ۵۲   | ۲۸   | ۱۶۶۴۱۶   | ۱۱۹۱۸   | ۲۰                                      |
| ۲۰                | ۷                         | ۲۸  | ۵۷                                  | ۵۳   | ۲۹   | ۱۳۶۲۹۸   | ۱۲۹۶۶   | ۲۰                                      |
|                   |                           | متابولیت‌های شبکه‌ی اصلی (۷۲)             | واکنش‌های شبکه‌ی اصلی (۹۴)          | واکنش‌های شبکه‌ی اصلی (۹۴)                         | شبکه‌ی اصلی (۲۸) ستون‌های کرنل در شبکه‌ی اصلی            | مدهای شبکه‌ی اصلی (۱۶۶۷۳)  | زمان شبکه‌ی اصلی (۱۸۵۴ میلی ثانیه)                      | پیمایش‌ها برای شبکه‌ی اصلی (۲۲)         |

این مدل، یکی از بزرگترین مدل‌هایی است که در حال حاضر محاسبه‌ی مدهای پایه‌ی آن بر روی کامپیوترهای شخصی ممکن است. البته این مدل بخشی از متابولیسم یک باکتری بوده و تا حد زیادی تغییر یافته است. برای مثال واکنش‌هایی در این شبکه وجود دارد که با تعداد زیادی متابولیت متصل هستند. این واکنش‌ها طبیعتاً واکنش‌های واقعی نمی‌باشند و احتمالاً برای اعمال تاثیر بقیه‌ی شبکه یا برای اعمال روش‌های بهینه‌سازی خطی به شبکه اضافه شده‌اند. با توجه به این جدول می‌توان گفت که برای این مدل در همه‌ی حالات اوضاع از لحاظ زمان اجرا بدتر شده است. در مورد این شبکه در بهترین حالت (دو زیرشبکه) تعداد متابولیت‌ها به ۲۱ و تعداد واکنش‌ها به ۴۴ عدد کاهش یافته است. تعداد ستون‌های ماتریس کرنل در بهترین حالت (۱۴ زیرشبکه) تنها دو عدد کاهش یافته است. ضمن این که با استفاده از روش‌های فشرده‌سازی موجود، تعداد متابولیت‌ها به ۶۰ و تعداد واکنش‌ها به ۸۰ کاهش یافته است.

آزمایش ۳ (*Mycoplasma pneumonia 309*) [۶۱]: آزمایش سوم بر روی یک مدل نسبتاً بزرگ (۵۴۹ متابولیت و ۱۰۸۹ واکنش) انجام گرفته است. در این آزمایش امکان محاسبه‌ی مدهای پایه بر روی کامپیوتر شخصی وجود نداشته است. بنابراین تنها اندازه‌ی شبکه و ماتریس کرنل بعد از درشت‌دانه‌سازی گزارش شده است (جدول ۳-۵).

جدول ۳-۵ نتایج آزمایش ۳

| تعداد زیر شبکه‌ها | تعداد متابولیت‌های ساختگی | تعداد واکنش‌ها پس از درشت‌دانه‌سازی | تعداد واکنش‌ها پس از درشت‌دانه‌سازی | تعداد واکنش‌ها پس از درشت‌دانه‌سازی | تعداد واکنش‌ها پس از درشت‌دانه‌سازی |
|-------------------|---------------------------|-------------------------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| ۲                 | ۰                         | ۱۳۱                                 | ۵۲۰                                 | ۴۳۷                                 | ۳۰۶                                 |
| ۳                 | ۰                         | ۱۳۱                                 | ۵۲۰                                 | ۴۳۷                                 | ۳۰۶                                 |
| ۴                 | ۰                         | ۱۳۱                                 | ۵۲۰                                 | ۴۴۰                                 | ۳۰۹                                 |
| ۵                 | ۰                         | ۱۳۱                                 | ۵۲۰                                 | ۴۳۸                                 | ۳۰۷                                 |
| ۶                 | ۰                         | ۱۳۱                                 | ۵۲۰                                 | ۴۳۹                                 | ۳۰۸                                 |
| ۷                 | ۰                         | ۱۳۱                                 | ۵۲۰                                 | ۴۴۱                                 | ۳۱۰                                 |
| ۸                 | ۰                         | ۱۳۱                                 | ۵۲۰                                 | ۴۴۰                                 | ۳۰۹                                 |
| ۹                 | ۰                         | ۱۳۱                                 | ۵۲۰                                 | ۴۴۰                                 | ۳۰۹                                 |
| ۱۰                | ۰                         | ۱۳۱                                 | ۵۲۰                                 | ۴۴۰                                 | ۳۰۹                                 |
| ۱۱                | ۰                         | ۱۳۱                                 | ۵۲۰                                 | ۴۴۱                                 | ۳۱۰                                 |
| ۱۲                | ۰                         | ۱۳۱                                 | ۵۲۰                                 | ۴۳۸                                 | ۳۰۷                                 |
| ۱۳                | ۰                         | ۱۳۱                                 | ۵۲۰                                 | ۴۳۸                                 | ۳۰۷                                 |
| ۱۴                | ۰                         | ۱۳۲                                 | ۵۲۱                                 | ۴۴۱                                 | ۳۰۹                                 |
| ۱۵                | ۱                         | ۱۳۲                                 | ۵۲۱                                 | ۴۴۲                                 | ۳۱۰                                 |
| ۱۶                | ۱                         | ۱۳۲                                 | ۵۲۱                                 | ۴۴۲                                 | ۳۱۰                                 |
| ۱۷                | ۱                         | ۱۳۲                                 | ۵۲۱                                 | ۴۴۲                                 | ۳۱۰                                 |
| ۱۸                | ۱                         | ۱۳۲                                 | ۵۲۰                                 | ۴۳۹                                 | ۳۰۸                                 |
| ۱۹                | ۱                         | ۱۳۲                                 | ۵۲۱                                 | ۴۴۲                                 | ۳۱۰                                 |
| ۲۰                | ۱                         | ۱۳۲                                 | ۵۲۰                                 | ۴۴۰                                 | ۳۰۹                                 |

این شبکه پس از اعمال روش‌های فشرده‌سازی موجود، به شبکه‌ای با ۴۰۴ متابولیت و ۹۳۱ واکنش و ماتریس کرنل با ۵۳۳ ستون تبدیل شده است. در حالی که پس از فشرده‌سازی توسط روش درشت‌دانه

سازی در بهترین حالت شبکه‌ای با ۱۳۱ متابولیت و ۴۳۷ واکنش و ماتریس کرنل با ۳۰۶ ستون به دست آمده است.

آزمایش ۴ (*Escherichia coli*\_BW2952 [۶۲]): آزمایش چهارم بر روی یک شبکه‌ی متابولیکی مقیاس ژنوم با اندازه‌ی بزرگ (۱۹۴۹ متابولیت و ۲۷۴۲ واکنش) انجام گرفته است. نتایج این آزمایش در جدول ۴-۵ گزارش شده است.

جدول ۴-۵ نتایج آزمایش ۴

| تعداد زیر شبکه‌ها | تعداد متابولیت‌های ساختگی | تعداد متابولیت‌ها پس از درشت‌دانه سازی | تعداد واکنش‌ها پس از درشت‌دانه سازی | موازی درشت‌واکنش‌های از یکی کردن تعداد واکنش‌ها پس موازی | ستون‌های کرنل پس از یکی کردن درشت‌واکنش‌های موازی |
|-------------------|---------------------------|--|-------------------------------------|--|---|
| ۲                 | ۴                         | ۳۱۷                                    | ۲۳۷۴                                | ۹۹۸  | ۶۸۶   |
| ۳                 | ۴                         | ۳۱۷                                    | ۲۰۲۸                                | ۹۸۲  | ۶۷۰   |
| ۴                 | ۶                         | ۳۱۹                                    | ۲۳۷۶                                | ۹۹۹  | ۶۸۵   |
| ۵                 | ۱۰                        | ۳۲۳                                    | ۲۰۳۲                                | ۹۸۶  | ۶۶۸   |
| ۶                 | ۱۱                        | ۳۲۴                                    | ۲۰۳۳                                | ۹۸۸  | ۶۶۹   |
| ۷                 | ۸                         | ۳۲۱                                    | ۲۰۳۰                                | ۹۸۳  | ۶۶۷   |
| ۸                 | ۱۵                        | ۳۲۸                                    | ۲۰۳۷                                | ۹۹۰  | ۶۶۷   |
| ۹                 | ۸                         | ۳۲۱                                    | ۲۰۲۹                                | ۹۸۲  | ۶۶۷   |
| ۱۰                | ۹                         | ۳۲۲                                    | ۲۰۳۱                                | ۹۸۶  | ۶۶۹   |
| ۱۱                | ۱۶                        | ۳۲۹                                    | ۲۰۳۸                                | ۹۹۱  | ۶۶۷   |
| ۱۲                | ۷                         | ۳۲۰                                    | ۲۰۳۱                                | ۹۸۵  | ۶۷۰   |
| ۱۳                | ۱۷                        | ۳۳۰                                    | ۲۰۳۹                                | ۹۹۴  | ۶۶۹   |
| ۱۴                | ۱۰                        | ۳۲۳                                    | ۲۰۳۲                                | ۹۸۶  | ۶۶۸   |
| ۱۵                | ۱۶                        | ۳۲۹                                    | ۲۰۳۸                                | ۹۹۱  | ۶۶۷   |
| ۱۶                | ۱۴                        | ۳۲۷                                    | ۲۰۳۶                                | ۹۸۹  | ۶۶۷   |
| ۱۷                | ۱۴                        | ۳۲۷                                    | ۲۰۳۶                                | ۹۹۱  | ۶۶۹   |
| ۱۸                | ۱۱                        | ۳۲۴                                    | ۲۰۳۳                                | ۹۸۸  | ۶۶۹   |
| ۱۹                | ۸                         | ۳۲۱                                    | ۲۰۳۰                                | ۹۸۳  | ۶۶۷   |
| ۱۱۴               | ۴۱                        | ۳۳۱                                    | ۲۰۴۰                                | ۹۵۱  | ۶۰۳   |

پس از اعمال روش‌های فشردسازی موجود، شبکه‌ای با ۱۰۷۳ متابولیت و ۱۸۰۹ واکنش و ماتریس کرنل با ۸۴۵ ستون حاصل می‌شود. در مقابل با اعمال روش درشت‌دانه‌سازی، شبکه‌ی حاصل در بهترین حالت (با ۱۱۴ زیرشبکه) دارای ۳۳۱ متابولیت، ۹۵۱ واکنش و ماتریس کرنل با ۶۰۳ ستون بوده است.

### ۵-۳ تحلیل نتایج

نتایج چهار آزمایش انجام شده حاکی از این هستند که روش درشت‌دانه‌سازی شبکه در بیشتر موارد اندازه‌ی شبکه را به طور قابل توجهی کاهش می‌دهد. در مورد شبکه‌های متابولیکی، روش‌های متعددی، موسوم به روش‌های فشردسازی یا کاهش شبکه، برای کاهش اندازه‌ی این شبکه‌ها قبل از محاسبه‌ی مدهای پایه انجام می‌گیرد. هدف این روش‌ها این است که ماتریس کرنل که عملیات محاسبه‌ی مدهای پایه بر روی آن آغاز می‌شود، دارای حداقل اندازه بوده و فاقد هرگونه افزونگی باشد [۱۳، ۱۴، ۲۴]. مقایسه‌ی این روش‌ها با روش درشت‌دانه‌سازی نشان می‌دهد که میزان فشردسازی با استفاده از روش درشت‌دانه‌سازی در مورد شبکه‌های مقیاس ژنوم، به مراتب از میزان فشردسازی روش‌های موجود بیشتر است. در جدول ۵-۵ مقایسه‌ی نتایج روش‌های فشردسازی موجود و روش درشت‌دانه‌سازی شبکه آورده شده است.

جدول ۵-۵ مقایسه‌ی نتایج روش‌های فشردسازی موجود با روش درشت‌دانه‌سازی شبکه

| آزمایش | روش‌های فشردسازی موجود |               | روش درشت‌دانه‌سازی |               |                           |
|--------|------------------------|---------------|--------------------|---------------|---------------------------|
|        | کاهش متابولیت‌ها       | کاهش واکنش‌ها | کاهش متابولیت‌ها   | کاهش واکنش‌ها | کاهش ستون‌های ماتریس کرنل |
| ۱      | ٪۶۴                    | ٪۵۸           | ٪۸۷                | ٪۶۵           | ٪-۳۸                      |
| ۲      | ۱۵٪                    | ۱۷٪           | ۷۱٪                | ۵۳٪           | ۴٪                        |
| ۳      | ۲۶٪                    | ۱۵٪           | ۷۶٪                | ٪۶۰           | ۴۳٪                       |
| ۴      | ۴۴٪                    | ۳۴٪           | ۸۳٪                | ٪۶۵           | ۲۹٪                       |

با توجه به نتایج مشاهده شده می‌توان گفت که عملکرد روش درشت‌دانه‌سازی بسیار وابسته به ماهیت شبکه است. که البته با توجه به این که ایده‌ی این روش بر مبنای وجود درجه‌ای از ماژولاریتی طراحی شده است، چنین مشاهده‌ای دور از انتظار نیز نمی‌باشد. البته نتایج نشان می‌دهد که اندازه‌ی شبکه در تقریباً همه موارد کاهش یافته است.

کاهش اندازه‌ی شبکه و ماتریس کرنل هدف اصلی همه‌ی روش‌های فشرده‌سازی بوده و این روش‌ها در همه‌ی ابزارهای محاسبه‌ی مدهای پایه قبل از شروع عملیات به کار می‌روند. زیرا انتظار می‌رود که با کاهش اندازه‌ی شبکه عموماً زمان اجرا نیز کاهش یابد. اما به هر حال به دلیل این که شناخت دقیقی از رابطه‌ی زمان اجرای الگوریتم پوچ‌فضا و پارامترهای شبکه وجود ندارد [۳۰، ۳۱]، نمی‌توان در مورد تاثیر بر روی زمان الگوریتم پوچ‌فضا اظهار نظر قطعی کرد. برای مثال در مورد شبکه‌ای مانند شبکه‌ی مربوط به آزمایش ۱ زمان اجرا حدود ۵۰٪ کاهش یافته اما در مورد شبکه‌ی مربوط به آزمایش ۲ زمان اجرا در همه‌ی حالات بدتر شده است. در مورد شبکه‌ی مربوط به آزمایش ۳ و ۴ با وجود این که اندازه‌ی شبکه به شکل قابل ملاحظه‌ای کاهش یافته اما هنوز به اندازه‌ی نرسیده است که بتوان مدهای پایه‌ی آن را بتوان با استفاده از کامپیوترهای شخصی محاسبه کرد. با توجه به این که اکنون محاسبه‌ی برخی از شبکه‌های مقیاس ژنوم نسبتاً کوچک‌تر، بر روی بسترهای پردازش موازی انجام گرفته [۲۴، ۲۷]، کاهش اندازه‌ی شبکه ممکن است بتواند پردازش شبکه‌های بزرگتری را بر روی این بسترها ممکن کند.

روش درشت‌دانه‌سازی همچنین می‌تواند با روش‌های موجود ترکیب شده و کاهش بیشتری را روی اندازه‌ی شبکه و ماتریس کرنل به وجود بیاورد. به علاوه این روش را می‌توان با روش‌های نمونه برداری ([۳۵-۳۳]) یا روش‌های دیگری که رویکرد محاسبه‌ی بخشی از مدهای پایه را دارند ([۳۹، ۳۸، ۳۲])، ترکیب کرد. برای مثال روش‌هایی که با استفاده از انرژی آزاد گیبس مسیرهای غیر ممکن را حذف می‌کنند ([۳۸]) به خوبی می‌توانند با این روش ترکیب گردند. این کار می‌تواند به این ترتیب صورت گیرد که با حذف درشت‌واکنش‌هایی که از لحاظ انرژی آزاد گیبس نمی‌توانند به طور خود به خودی روی دهند، اندازه‌ی شبکه‌ی درشت‌دانه را کوچک‌تر کرد.

از مزایای دیگر روش درشت‌دانه‌سازی این است که شبکه‌ی درشت‌دانه همواره از لحاظ شهودی معنی‌دار می‌باشد. یعنی درشت‌واکنش‌ها درست مانند واکنش‌های عادی به طور معنی‌داری تبدیل مواد را به یکدیگر نشان می‌دهند. این مزیت نه تنها باعث می‌گردد که مدهای پایه‌ی شبکه‌ی درشت‌دانه را بتوان

بدون برگرداندن آن‌ها به شبکه‌ی اصلی مورد بررسی قرار داد، بلکه شبکه‌ی درشت‌دانه می‌تواند جنبه‌هایی از شبکه را نشان دهد که به راحتی نمی‌توان داخل شبکه‌ی اصلی دید. برای مثال در مورد شبکه‌ی مربوط به آزمایش ۴، تعداد زیادی از درشت‌واکنش‌های به دست آمده با یکدیگر موازی هستند به‌طوری‌که پس از یکی کردن درشت‌واکنش‌های موازی تعداد درشت‌واکنش‌ها به حدود نصف کاهش می‌یابد. درشت‌واکنش‌های موازی درواقع نشان‌دهنده‌ی انجام یک فعالیت متابویکی از مسیرهای مختلف (مسیرهای افزونه) هستند. بنابراین شبکه‌ی درشت‌دانه خود می‌تواند وسیله‌ای برای کسب اطلاعات شهودی و کمی باشد که از شبکه‌ی اصلی قابل برداشت نمی‌باشد.

۶

## فصل ششم

### جمع‌بندی و کارهای آینده



## جمع‌بندی و کارهای آینده

در این پایان‌نامه ابتدا با بیان چالش‌های استفاده از روش مدهای پایه در شبکه‌های متابولیکی مقیاس ژنوم به بیان تاریخچه‌ی موضوع و تلاش‌های مختلفی که در رویارویی با این چالش‌ها صورت گرفته است پرداخته شد. سپس راهکاری بر اساس تجزیه‌ی شبکه و درشت‌دانه‌سازی واکنش‌های شبکه معرفی شد. این روش به منظور تبدیل مساله به مساله‌های کوچکتر انجام می‌گیرد. نتایج اجرای این روش روی یک شبکه‌ی نمونه باعث کاهش اجرا شده است. برای شبکه‌های بزرگ همچنان امکان محاسبه‌ی مدهای پایه بر روی کامپیوترهای عادی وجود ندارد. اما شاخص اصلی پیچیدگی مساله یعنی اندازه (تعداد ستون‌های) ماتریس کرنل برای بخش قابل توجهی از شبکه‌های موجود کاهش می‌یابد. بنابراین هرچند همچنان نمی‌توان مدهای پایه‌ی این شبکه‌ها را با استفاده از کامپیوترهای عادی محاسبه کرد، اما قاعده‌تاً کاهش پیچیدگی مساله می‌تواند برای به دست آوردن مدهای پایه‌ی شبکه‌های مقیاس ژنوم بر روی بسترهای محاسبات با کارایی بالا کمک کننده باشد.

در این روش طبق یک چارچوب خاص، شبکه به زیرشبکه‌هایی تجزیه می‌شود. یکی از این شبکه‌ها به عنوان شبکه‌ی مرکزی نقش ارتباط بین زیرشبکه‌های دیگر را دارد. تجزیه با این معیار انجام می‌گردد که تا جای ممکن زیرشبکه‌ها از یکدیگر جدا بوده و تنها دارای اتصالات داخلی یا اتصالاتی با شبکه‌ی مرکزی باشند. ضمن این که سعی می‌شود شبکه‌ی مرکزی تا حد امکان کوچک باشد. با به دست آوردن چنین تجزیه‌ای، شبکه‌ی متابولیکی داده‌شده به یک شبکه‌ی درشت‌دانه تبدیل می‌گردد. سپس از مدهای پایه‌ی شبکه‌ی درشت‌دانه مدهای پایه‌ی شبکه‌ی اصلی به دست می‌آیند. این روش هنگامی سودمند است که شبکه‌ی درشت‌دانه دارای ابعاد و پیچیدگی کمتری نسبت به شبکه‌ی اصلی باشد. سربار اضافه شده شامل عملیات تجزیه، محاسبه‌ی مدهای پایه‌ی زیرشبکه‌ها، و حذف بردارهای غیر پایه از جواب‌های به دست آمده است.

برای بهبود کارایی این روش می‌توان به ایده‌های متعددی اشاره کرد. برای مثال در روش ارائه شده در این تحقیق برای نشان دادن محدودیت‌هایی که یک زیر شبکه بر روی مرزهای خود ایجاد می‌کند، از مدهای پایه‌ی آن زیرشبکه استفاده شده است. اما محدودیت‌های مرزی یک شبکه‌ی متابولیکی را می‌توان با استفاده از مدهای مولد شار نیز نشان داد که تعداد آن‌ها معمولاً کمتر از مدهای پایه‌ی شار

است. بنابراین با استفاده از این روش می‌توان شبکه‌ی درشت‌دانه کوچکتری به دست آورد. چالش چنین رویکردی این است که پس از به دست آوردن مسیرهایی که شامل مدهای مولدشار هستند، تبدیل این مسیرهایی برحسب مدهای پایه ممکن است ساده نباشد. ضمن این که رابطه‌ی مسیرهایی که از این راه به دست می‌آیند با مجموعه‌ی مدهای پایه مشخص نیست. یک رویکرد دیگر می‌تواند این باشد که تجزیه‌ی شبکه را با همین چارچوب منتها با روش دیگر یا با معیار دیگری انجام داد. برای مثال در آزمایش‌های انجام شده در این تحقیق، متابولیت‌های میانی با استفاده از تعریف یک آستانه‌ی ۴ برای درجه‌ی رئوس تعریف می‌شوند. در حالی که ممکن است بتوان با روش بهتری چنین متابولیت‌هایی را مشخص کرد. همچنین در تجزیه‌ی زیرشبکه‌ها به جای معیار کاهش اندازه‌ی برش می‌توان از معیار دیگری استفاده کرد. با توجه به این که یکی شدن واکنش‌های موازی تاثیر قابل توجهی در کاهش تعداد واکنش‌ها و اندازه‌ی ماتریس کرنل دارند. بنابراین اگر بتوان معیاری پیدا کرد که تجزیه‌ی شبکه را با هدف افزایش تعداد واکنش‌های موازی (به جای کاهش اندازه‌ی برش) انجام داد ممکن است بتوان نتایج بهتری به دست آورد. همچنین به دلیل سادگی و شهودی بودن روش ممکن است بتوان این روش را به طور معنی‌داری با روش‌هایی که قصد محاسبه‌ی زیرمجموعه‌ای از مدهای پایه را دارند ترکیب نمود.

## منابع و مراجع

- [1] C. Wagner and R. Urbanczik, "The geometry of the flux cone of a metabolic network," *Biophysical Journal*, vol. 89, no. 6, pp. 3837-3845, 2005.
- [2] C. Francke, R. J. Siezen, and B. Teusink, "Reconstructing the metabolic network of a bacterium from its genome," *Trends in Microbiology*, vol. 13, no. 11, pp. 550-558, 2005.
- [3] C. S. Henry, M. DeJongh, A. A. Best, P. M. Frybarger, B. Linsay, and R. L. Stevens, "High-throughput generation, optimization and analysis of genome-scale metabolic models," *Nature Biotechnology*, vol. 28, no. 9, pp. 977-982, 2010.
- [4] Z. A. King *et al.*, "BiGG Models: A platform for integrating, standardizing and sharing genome-scale models," *Nucleic Acids Research*, vol. 44, no. D1, pp. D515-D522, 2016.
- [5] V. Lacroix, L. Cottret, P. Thébault, and M.-F. Sagot, "An introduction to metabolic networks and their structural analysis," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 5, no. 4, pp. 594-617, 2008.
- [6] P. Holme and M. Huss, "Substance graphs are optimal simple-graph representations of metabolism," *Chinese Science Bulletin*, vol. 55, no. 27-28, pp. 3161-3168, 2010.
- [7] E. Ullah, S. Aeron, and S. Hassoun, "gEFM: an algorithm for computing elementary flux modes using graph traversal," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 13, no. 1, pp. 122-134, 2016.
- [8] R. U. Ibarra, J. S. Edwards, and B. O. Palsson, "Escherichia coli K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth," *Nature*, vol. 420, no. 6912, pp. 186-189, 2002.
- [9] A. P. Burgard, E. V. Nikolaev, C. H. Schilling, and C. D. Maranas, "Flux coupling analysis of genome-scale metabolic network reconstructions," *Genome Research*, vol. 14, no. 2, pp. 301-312, 2004.
- [10] S. A. Marashi and A. Bockmayr, "Flux coupling analysis of metabolic networks is sensitive to missing reactions," *Biosystems*, vol. 103, no. 1, pp. 57-66, Jan 2011.
- [11] S.-A. Marashi, L. David, and A. Bockmayr, "On flux coupling analysis of metabolic subsystems," *Journal of Theoretical Biology*, vol. 302, pp. 62-69, 2012.
- [12] A. C. Muller and A. Bockmayr, "Fast thermodynamically constrained flux variability analysis," *Bioinformatics*, vol. 29, no. 7, pp. 903-9, Apr 01 2013.

- [13] J. Gagneur and S. Klamt, "Computation of elementary modes: a unifying framework and the new binary approach," *BMC Bioinformatics*, vol. 5, p. 175, Nov 04 2004.
- [14] S. Klamt, J. Gagneur, and A. von Kamp, "Algorithmic approaches for computing elementary modes in large biochemical reaction networks," presented at the IEE Proceedings-Systems Biology, 2005.
- [15] A. von Kamp and S. Schuster, "Metatool 5.0: fast and flexible elementary modes analysis," *Bioinformatics*, vol. 22, no. 15, pp. 1930-1, Aug 01 2006.
- [16] S. Schuster and C. Hilgetag, "On Elementary Flux Modes in biochemical reaction systems at steady state," *Journal of Biological Systems*, vol. 02, no. 02, pp. 165-182, 1994.
- [17] S. Schuster, D. A. Fell, and T. Dandekar, "A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks," *Nature Biotechnology*, vol. 18, no. 3, pp. 326-332, 2000.
- [18] E. Fischer and U. Sauer, "A novel metabolic cycle catalyzes glucose oxidation and anaplerosis in hungry *Escherichia coli*," *Journal of Biological Chemistry*, vol. 278, no. 47, pp. 46446-46451, 2003.
- [19] J. Zanghellini, D. E. Ruckerbauer, M. Hanscho, and C. Jungreuthmayer, "Elementary flux modes in a nutshell: properties, calculation and applications," *Biotechnology Journal*, vol. 8, no. 9, pp. 1009-1016, 2013.
- [20] A. Rezola, J. Pey, L. Tobalina, A. Rubio, J. E. Beasley, and F. J. Planes, "Advances in network-based metabolic pathway analysis and gene expression data integration," *Briefings in Bioinformatics*, vol. 16, no. 2, pp. 265-79, Mar 2015.
- [21] C. Wagner, "Nullspace approach to determine the elementary modes of chemical reaction systems," *The Journal of Physical Chemistry B*, vol. 108, no. 7, pp. 2425-2431, 2004.
- [22] R. Urbanczik and C. Wagner, "An improved algorithm for stoichiometric network analysis: theory and applications," *Bioinformatics*, vol. 21, no. 7, pp. 1203-1210, 2005.
- [23] T. Pfeiffer, I. Sanchez-Valdenebro, J. C. Nuno, F. Montero, and S. Schuster, "METATOOL: for studying metabolic networks," *Bioinformatics*, vol. 15, no. 3, pp. 251-7, Mar 1999.
- [24] D. Jevremović, C. T. Trinh, F. Sreenc, C. P. Sosa, and D. Boley, "Parallelization of nullspace algorithm for the computation of metabolic pathways," *Parallel Computing*, vol. 37, no. 6, pp. 261-278, 2011.
- [25] M. Terzer and J. Stelling, "Large-scale computation of elementary flux modes with bit pattern trees," *Bioinformatics*, vol. 24, no. 19, pp. 2229-2235, 2008.
- [26] D. Jevremovic, D. Boley, and C. P. Sosa, "Divide-and-conquer approach to the parallel computation of elementary flux modes in metabolic networks," presented at the IEEE International Symposium on Parallel and Distributed Processing Workshops and Phd Forum, 2011.

- [27] K. A. Hunt, J. P. Folsom, R. L. Taffs, and R. P. Carlson, "Complete enumeration of elementary flux modes through scalable demand-based subnetwork definition," *Bioinformatics*, vol. 30, no. 11, pp. 1569-78, Jun 01 2014.
- [28] J. F. H. Céspedes, F. D. A. G. Asensio, and J. M. G. Carrasco, "A new approach to obtain efms using graph methods based on the shortest path between end nodes," in *International Conference on Bioinformatics and Biomedical Engineering*, 2015, pp. 641-649: Springer.
- [29] E. Ullah, K. Lee, and S. Hassoun, "An algorithm for identifying dominant-edge metabolic pathways," in *Computer-Aided Design-Digest of Technical Papers, 2009. ICCAD 2009. IEEE/ACM International Conference on*, 2009, pp. 144-150: IEEE.
- [30] V. Acuna, F. Chierichetti, V. Lacroix, A. Marchetti-Spaccamela, M. F. Sagot, and L. Stougie, "Modes and cuts in metabolic networks: complexity and algorithms," *Biosystems*, vol. 95, no. 1, pp. 51-60, Jan 2009.
- [31] V. Acuna, A. Marchetti-Spaccamela, M. F. Sagot, and L. Stougie, "A note on the complexity of finding and enumerating elementary modes," *Biosystems*, vol. 99, no. 3, pp. 210-4, Mar 2010.
- [32] L. F. de Figueiredo *et al.*, "Computing the shortest elementary flux modes in genome-scale metabolic networks," *Bioinformatics*, vol. 25, no. 23, pp. 3158-65, Dec 01 2009.
- [33] D. Machado, Z. Soons, K. R. Patil, E. C. Ferreira, and I. Rocha, "Random sampling of elementary flux modes in large-scale metabolic networks," *Bioinformatics*, vol. 28, no. 18, pp. i515-i521, Sep 15 2012.
- [34] C. Kaleta, L. s. F. De Figueiredo, J. Behre, and S. Schuster, "EFMEvolver: Computing elementary flux modes in genome-scale metabolic networks," presented at the 14th German Conference on Bioinformatics, 2009.
- [35] J. Pey *et al.*, "TreeEFM: calculating elementary flux modes using linear optimization in a tree-based algorithm," *Bioinformatics*, p. btu733, 2014.
- [36] J. Pey and F. J. Planes, "Direct calculation of elementary flux modes satisfying several biological constraints in genome-scale metabolic networks," *Bioinformatics*, p. btu193, 2014.
- [37] L. David and A. Bockmayr, "Computing elementary flux modes involving a set of target reactions," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 11, no. 6, pp. 1099-1107, 2014.
- [38] M. P. Gerstl, D. E. Ruckerbauer, D. Mattanovich, C. Jungreuthmayer, and J. Zanghellini, "Metabolomics integrated elementary flux mode analysis in large metabolic networks," *Scientific Reports*, vol. 5, p. 8930, 2015.
- [39] M. P. Gerstl, C. Jungreuthmayer, and J. Zanghellini, "tEFMA: computing thermodynamically feasible elementary flux modes in metabolic networks," *Bioinformatics*, vol. 31, no. 13, pp. 2232-4, Jul 01 2015.

- [40] C. Jungreuthmayer, D. E. Ruckerbauer, and J. Zanghellini, "regEfmtool: speeding up elementary flux mode calculation using transcriptional regulatory rules in the form of three-state logic," *Biosystems*, vol. 113, no. 1, pp. 37-9, Jul 2013.
- [41] A. Rezola *et al.*, "Exploring metabolic pathways in genome-scale networks via generating flux modes," *Bioinformatics*, vol. 27, no. 4, pp. 534-540, 2011.
- [42] A. Larhlimi and A. Bockmayr, "A new constraint-based description of the steady-state flux cone of metabolic networks," *Discrete Applied Mathematics*, vol. 157, no. 10, pp. 2257-2266, 2009.
- [43] R. M. Jungers, F. Zamorano, V. D. Blondel, A. V. Wouwer, and G. Bastin, "Fast computation of minimal elementary decompositions of metabolic flux vectors," *Automatica*, vol. 47, no. 6, pp. 1255-1259, 2011.
- [44] Z. I. T. A. Soons, E. C. Ferreira, and I. Rocha, "Identification of minimal metabolic pathway models consistent with phenotypic data," *Journal of Process Control*, vol. 21, no. 10, pp. 1483-1492, 2011.
- [45] S. M. Kelk, B. G. Olivier, L. Stougie, and F. J. Bruggeman, "Optimal flux spaces of genome-scale stoichiometric models are determined by a few subnetworks," *Scientific Reports*, vol. 2, p. 580, 2012.
- [46] A. C. Müller and A. Bockmayr, "Flux modules in metabolic networks," *Journal of Mathematical Biology*, vol. 69, no. 5, pp. 1151-1179, 2012.
- [47] J. A. Papin, J. L. Reed, and B. O. Palsson, "Hierarchical thinking in network biology: the unbiased modularization of biochemical networks," *Trends in Biochemical Sciences*, vol. 29, no. 12, pp. 641-647, 2004.
- [48] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A. L. Barabasi, "The large-scale organization of metabolic networks," *Nature*, vol. 407, no. 6804, pp. 651-4, Oct 05 2000.
- [49] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A. L. Barabasi, "Hierarchical organization of modularity in metabolic networks," *Science*, vol. 297, no. 5586, pp. 1551-5, Aug 30 2002.
- [50] S. Schuster, T. Pfeiffer, F. Moldenhauer, I. Koch, and T. Dandekar, "Exploring the pathway structure of metabolism: decomposition into subnetworks and application to *Mycoplasma pneumoniae*," *Bioinformatics*, vol. 18, no. 2, pp. 351-361, 2002.
- [51] W. S. Verwoerd, "A new computational method to split large biochemical networks into coherent subnets," *BMC Systems Biology*, vol. 5, no. 1, p. 25, 2011.
- [52] W. S. Verwoerd, "Interactive extraction of metabolic subnets – the Netsplitter software implementation," *Journal of Molecular Engineering and Systems Biology*, vol. 1, no. 1, p. 2, 2012.
- [53] H. W. Ma, X. M. Zhao, Y. J. Yuan, and A. P. Zeng, "Decomposition of metabolic network into functional modules based on the global connectivity structure of reaction graph," *Bioinformatics*, vol. 20, no. 12, pp. 1870-6, Aug 12 2004.

- [54] R. Guimera and L. A. Nunes Amaral, "Functional cartography of complex metabolic networks," *Nature*, vol. 433, no. 7028, pp. 895-900, Feb 24 2005.
- [55] A. C. Reimers, "Hierarchical decomposition of metabolic networks using k-modules," *Biochemical Society Transactions*, vol. 43, no. 6, pp. 1146-1150, 2015.
- [56] C. Kaleta, L. F. de Figueiredo, and S. Schuster, "Can the whole be less than the sum of its parts? Pathway analysis in genome-scale metabolic networks using elementary flux patterns," *Genome Research*, vol. 19, no. 10, pp. 1872-1883, 2009.
- [57] A. C. Reimers, F. J. Bruggeman, B. G. Olivier, and L. Stougie, "Fast flux module detection using matroid theory," *Journal of Computational Biology*, vol. 22, no. 5, pp. 414-424, 2015.
- [58] G. Karypis and V. Kumar, "Multilevel k-way hypergraph partitioning," *Vlsi Design*, vol. 11, no. 3, pp. 285-300, 2000.
- [59] E. J. Kerkhoven *et al.*, "Handling uncertainty in dynamic models: the pentose phosphate pathway in trypanosoma brucei," *PLOS Computational Biology*, vol. 9, no. 12, p. e1003371, 2013.
- [60] J. D. Orth, R. M. Fleming, and B. Ø. Palsson, "Reconstruction and use of microbial metabolic networks: the core Escherichia coli metabolic model as an educational guide," *EcoSal Plus*, vol. 4, no. 1, 2010.
- [61] T. Kenri *et al.*, "Complete genome sequence of Mycoplasma pneumoniae type 2a strain 309, isolated in Japan," *Journal of Bacteriology*, vol. 194, no. 5, pp. 1253-1254, 2012.
- [62] J. M. Monk *et al.*, "Genome-scale metabolic reconstructions of multiple Escherichia coli strains highlight strain-specific adaptations to nutritional environments," *Proceedings of the National Academy of Sciences*, vol. 110, no. 50, pp. 20338-20343, 2013.

## Abstract

Elementary Flux Modes (EFMs) have a great importance both as a mathematical definition for the concept of metabolic pathways and as a tool in metabolic engineering applications. Therefore, many researches are working on efficient computation of EFMs. However, their computation on networks with more than hundreds of reactions is impractical, even on large computing platforms. The idea of decomposing these networks to smaller subnetworks is pursued by number of researches, but these researches study the subnetworks in isolation and have not provided a way to connect these subnetwork EFMs. In this research, with definition of coarse reactions, the relationship between subnetwork efms and full network efms have been established. In this method, the problem of connecting subnetwork EFMs leads to a coarsened network. By computing the EFMs of the coarsened network, the efms of the original network can be obtained. Therefore the problem of computing the EFMs of a network is reduced to computing EFMs for a coarsened, usually smaller, network. The coarsening of the network in this way can be used as a network compression method. The result of this method, applied to 4 real metabolic models, indicate that it reduces the network size by more than 50%. Also, in large metabolic networks, the size of the kernel matrix, an important indicator of the running time of EFM computation algorithms, is reduced by 30-40%. Furthermore, for a medium size network, the runtime of the EFM computation is reduced by around 50% which proves that the network coarsening method presented here can reduce the time of the EFM computation. The results show that in all cases the network coarsening method reduces the number of reactions and metabolites more than existing methods.

**Key Words:** Metabolic Networks, Network Compression, Network Decomposition, Elementary Flux Modes, Coarse Reactions





**Amirkabir University of Technology**  
**(Tehran Polytechnic)**

**Computer and Information Technology Engineering Department**

**MSc Thesis**

**Title**

**Hierarchical Decomposition of Metabolic Networks for Calculation  
of Elementary Flux Modes**

**By**

**Sina Ghadermarzi**

**Supervisor**

**Dr. Morteza Saheb Zamani**

**Advisor**

**Dr. Seyed Amir Marashi**

**February 2017**