

## استفاده از روش دسته‌بندی طیفی سلسله مراتبی در شبکه‌های قدرت تحت وب

سید حسین رضوی نیا<sup>۱\*</sup>، دکتر ایمان پورفر<sup>۲</sup>

\*۱- دانشجوی ارشد، مهندسی برق- قدرت، دانشگاه آزاد دزفول، razavi1k@gmail.com

۲- استادیار، مهندسی برق، دانشگاه جندی شاپور، دزفول

### چکیده

سیستم انتقال برق را می‌توان با شبکه‌ای نشان داد که دارای گره و اتصالات می‌باشد که نشان دهنده شین‌ها و خطوط انتقال الکتریکی هستند. به هر خط می‌توان مقدار داد که نشان دهنده برخی از ویژگی‌های الکتریکی خط مثل ادمیتانس خط یا توان عبوری متوسط در زمان مشخص می‌باشد. از روش دسته‌بندی طیفی سلسله مراتبی برای آشکار سازی ساختار اتصالات داخلی یک شبکه استفاده می‌کنیم. دسته‌بندی طیفی از مقادیر و بردارهای ویژه ماتریس شبکه استفاده می‌کند، این ماتریس از نظر ریاضی پر کاربرد می‌باشد و برای هر نوع شبکه‌ای می‌توان از آن استفاده کرد. هنگام استفاده از ادمیتانس‌های خط این ماتریس ساختار اتصالات داخلی ثابت شبکه اصلی را نشان می‌دهد، در حالی که استفاده از توان عبوری نواحی مجزایی را نشان می‌دهد که کم‌ترین میزان قطع توان عبوری را داشته و بنابراین به جدا سازی کنترل شده نواحی مربوط می‌شود. روش پیشنهادی فراتر از روش استاندارد *k-mean* است و ساختار شبکه را به صورت کامل در غالب دیاگرام درختی نشان می‌دهد. برای دسته‌بندی طیفی سیستم‌های قدرت استدلال نظری آورده شد و نتایج یک سیستم آزمایشی کوچک با استفاده از این روش ارائه شد. یک توجیه نظری کامل از استفاده از الگوریتم دسته‌بندی طیفی سلسله مراتبی ارائه شد و در نهایت یک برنامه‌ی تحت وب جهت اجرای این روش در شبکه‌های قدرت طراحی شد و شبکه قدرت ۳۹ شینه *IEEE* در این نرم افزار پیاده سازی شد و نتایج حاصل از این نرم افزار در *MATLAB* نشان داده شد. از زبان اسکریپت نویسی *PHP* جهت طراحی این نرم افزار استفاده شد.

### کلیدواژگان

دسته بندی طیفی، جزیره‌ای شدن کنترل شده، نظریه گراف، فاصله الکتریکی، محاسبه و تحلیل سیستم قدرت

### ۱- معرفی

ایده تجزیه (تقسیم بندی، قسمت کردن، جدا کردن، دسته‌بندی) یک شبکه قدرت و تبدیل آن به قسمت‌های کوچکتر به مفهوم روش تقسیم بندی مربوط می‌شود که در دهه ۵۰ توسط G. Kron ارائه شد [۱] و بسیاری از افراد این ایده را دنبال کردند [۲] - [۴]. انگیزه اولیه حافظه کم و سرعت محاسبه کامپیوترهای ابتدایی بود. بعدها محاسبه موازی سبب اقدامات پژوهشی بسیار با هدف تقسیم سیستم قدرت به بخش‌های کوچکتر شد تا محاسبات به صورت موازی انجام شوند. تجزیه شبکه‌های بزرگ و بهم پیوسته و تبدیل آن به نواحی مجزایی که می‌توان آنها را ساده تر مدیریت کرد معمولاً با مفهوم اتصال الکتریکی<sup>۱</sup> که همان فاصله الکتریکی بین گره‌های شبکه برای تعریف نواحی شبکه است انجام می‌شود [۵]. به طور کلی تجزیه<sup>۲</sup> شبکه می‌تواند با استفاده از مفهوم شبکه‌های هوشمند<sup>۳</sup>، کنترل سیستم قدرت را انعطاف پذیر، توزیع شده و سازگار کند [۶]. تجزیه شبکه در بخش‌های مجزا و کنترل شده

برای جلوگیری از خاموشی<sup>۴</sup> مناطق وسیع مورد استفاده قرار می‌گیرد. رایج ترین روش پیشنهادی داشتن ژنراتورهای کند<sup>۵</sup> و منسجم برای عدم پایداری گذرا<sup>۶</sup> [۷] و روش‌های مبتنی بر جست و جوی گراف<sup>۷</sup> می‌باشد [۸] - [۹]. اخیراً روش تازه‌ای به نام دسته‌بندی طیفی<sup>۸</sup> برای جزیره‌ای شدن کنترل شده پیشنهاد شده است که مبتنی بر پیشرفت‌های اخیر در نظریه گراف می‌باشد [۱۰] - [۱۱]. روش دسته‌بندی طیفی سلسله مراتبی نشان می‌دهد ساختار داخلی شبکه معین با توجه به انتخاب هر پارامتر الکتریکی می‌تواند به یک خط انتقال مثل ادمیتانس خط<sup>۹</sup> و یا توان عبوری<sup>۱۰</sup> مربوط می‌شود. هم چنین می‌تواند چند دسته را به عنوان ورودی در نظر بگیرد یا این عدد را می‌توان به صورت مستقیم از داده‌ها محاسبه کرد. خروجی دسته‌بندی سلسله مراتبی<sup>۱۱</sup>

<sup>4</sup> Blackouts

<sup>5</sup> Slow-Coherent generators

<sup>6</sup> Transient stability

<sup>7</sup> Graph-Search

<sup>8</sup> Spectral clustering

<sup>9</sup> Line admittance

<sup>10</sup> Power flow

<sup>11</sup> Hierarchical clustering

<sup>1</sup> Electrical cohesiveness

<sup>2</sup> Splitting

<sup>3</sup> Smart grids

کنید وزن یال نباید منفی ( $w_{ij} \geq 0$ ) باشد و همچنین وزن یال باید متقارن ( $w_{ij} = w_{ji}$ ) باشد تا معادله لاپلاسین حاصل گردد (بخش ۲-ب را ببینید).

۳. وزن در شبکه های قدرت: برای مطالعه ساختار کاری شبکه قدرت باید از توابع وزن یال زیر استفاده نماییم:

- مکان شناسی:  $w_{ij} = 1$  برای همه  $(i, j) \in E$  این اتصال خالص شبکه<sup>۵</sup> را می‌سند.
- ادمیتانس<sup>۶</sup>:  $w_{ij} = Y_{ij} = 1 / |R_{ij} + jX_{ij}|$  که  $R_{ij}$  و  $X_{ij}$  به ترتیب مقاومت خطی<sup>۷</sup> و راکتانس<sup>۸</sup> هستند، این استحکام اتصالات (فاصله الکتریکی<sup>۹</sup>) را می‌سند. در این مقاله ما از مدل شبکه DC استفاده خواهیم کرد که در آن وزن‌ها عبارتند از  $w_{ij} = 1 / X_{ij}$ .
- توان عبوری متوسط<sup>۱۰</sup>:  $w_{ij} = (|P_{ij}| + |P_{ji}|) / 2$  که  $P_{ij}$  توان (حقیقی<sup>۱۱</sup>) از  $i$  به  $j$  است (اگر شبکه بدون تلف باشد داریم  $w_{ij} = |P_{ij}| = |P_{ji}|$ ) این وزن اهمیت خط در شرایط کاری را می‌سند. بار کم به این معناست که احتمالاً خط از دسته‌بندی جدا خواهد شد.

توجه کنید که دو مقدار (وزن) اول ثابت هستند، یعنی برای یک سیستم قدرت ثابت می‌باشند، اما توان عبوری پویا و متغیر است و بر اساس شرایط کاری واقعی تغییر می‌کند.

وزن یال‌ها را می‌توان به عنوان معیاری برای قطع کردن آن خط در هنگام دسته‌بندی در نظر گرفت، با این حال معیار استحکام اتصال گره‌ها به هم برای قرار گرفتن در یک دسته‌بندی نیز می‌باشد. بنابراین ادمیتانس بر اساس دسته‌بندی، ساختار داخلی<sup>۱۲</sup> (فاصله الکتریکی) شبکه را نشان خواهد داد در حالی که توان عبوری بر اساس دسته‌بندی جزایری<sup>۱۳</sup> را نشان می‌دهد که وقتی جدا شوند حداقل یک توان عبوری در شبکه قطع می‌شود و در نتیجه می‌توانند برای جزیره‌ای کردن پیشگیرانه مفید باشند [۱۱].

ب. گراف لاپلاسین

ماتریس لاپلاسین در نظریه گراف در مهندس برق کاربرد گسترده‌ای دارد. دو نوع ماتریس لاپلاسین وجود دارند که به گراف ساده و بدون جهت و مقدار  $G = (V, E, w)$  مربوط می‌شوند.

۱. لاپلاسین نرمالیزه نشده: لاپلاسین  $G$  یک ماتریس  $N \times N$  است که به صورت زیر تعریف می‌گردد:

$$[L]_{ij} = \begin{cases} d_i & \text{if } i=j \\ -w_{ij} & \text{if } i \neq j \text{ and } (i,j) \in E \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

این یک ماتریس حقیقی متقارن<sup>۱۴</sup> است که به جز قطر اصلی، نامشبت دارد و مجموع هر ستون (یا ردیف) صفر است.

شبکه بر اساس اتصالات است. این را می‌توان به عنوان تجزیه کارکردی سیستم و تبدیل آن به سیستم‌های فرعی کوچکتر در شین‌های متصل در نظر گرفت. روش پیشنهادی با تجزیه شبکه و تبدیل آن به چند دسته، فراتر از تکنیک‌های دو بخشی یا دو بخشی بازگشتی<sup>۱</sup> است. در واقع به جای یک بخش از شبکه، تمام زیر ساخت‌های سلسله مراتبی بنیادی را با استفاده از الگوریتم دسته‌بندی سلسله مراتبی به عنوان گام نهایی نگه داشته می‌شود. بدین ترتیب می‌توان به صورت همزمان سطوح مختلف در دسته‌بندی و وابستگی داده‌های حاصل به مقیاس را نمایش داد که خود مبین سلسله مراتب کاری شبکه است. این تحقیق را باید اثباتی برای مفهوم روش دسته‌بندی طیفی در حوزه مهندسی برق در نظر گرفت [۱۲]-[۱۳] و در توضیح دادن اساس ریاضی آن دقت زیادی به خرج داد. پتانسیل این روش با جزئیات کامل در یک شبکه آزمایشی کوچک و شبکه‌های بزرگتر نشان داده شد. این مقاله نتیجه ترکیب مهندسی برق و ریاضیات است و می‌دانیم که اطلاعات با ارزشی در نظریه عملگر و گراف وجود دارند که در جامعه مهندسی برق خیلی شناخته نیستند ولی می‌توانند برای حل مسائل عملی مهندسی برق مفید واقع شوند. این مقاله در این جهت می‌کوشد.

## ۲- مقدمه

### الف. نمایش نموداری شبکه انتقال برق

۱- شرایط و نکات لازم: یک شبکه قدرت را می‌توان مانند یک نمودار نشان داد: راس (گره‌ها) نشان دهنده شین‌ها و یال‌ها (اتصالات) مبین اتصالات الکتریکی<sup>۲</sup> هستند. برای نمودار با رئوس  $V$  و یال‌های  $E$  نوشته می‌شود  $G=(V, E)$ . در ادامه فقط گراف‌های ساده در نظر گرفته می‌شود که هیچ حلقه و یال چند تایی ندارند. این فرض عمومیت ملاحظات را محدود نمی‌کند و می‌توان یال‌های چند تایی را با یک یال معادل جایگزین کرد. چون گراف  $G$  متناهی و ساده است می‌توان نوشت  $V = \{1, 2, \dots, N\}$  که  $N$  تعداد رئوس (گره‌ها) و  $E \subset V \times V$  که  $(i, j) \in E$  نشان دهنده یک یال (خط انتقال یا ترانسفورماتور) از گره  $i$  به گره  $j$  می‌باشد. چون دسته‌بندی طیفی جهت گره‌ها را نادیده می‌گیرد فرض می‌شود که تمام گراف‌ها بدون جهت هستند:  $(i, j) \in E$  اگر و تنها اگر  $(j, i) \in E$ .

۲- وزن گره‌ها: ساختار مکان شناسانه (توپولوژیکی<sup>۳</sup>) گراف در مورد شبکه قدرت اطلاعاتی در اختیار ما قرار نمی‌دهد. برای افزودن این اطلاعات از مقادیر گره‌ها استفاده می‌کنیم. مقدار گره تابعی از  $R \geq 0$   $V \times V \rightarrow R$  است به طوری که:

۱.  $w(i, j) = w(j, i)$  برای همه  $i, j$
۲.  $w(i, j) = 0$  اگر  $(i, j) \notin E$
۳.  $w(i, j) = \sum_{i=1}^N w(i, j)$

می‌باشد. اگر  $i \neq j$  و  $d_i = w(i, i)$  برای "مقدار گره دارای وزن"<sup>۴</sup> از نوتاسیون  $w_{ij} = w(i, j)$  استفاده می‌کنیم. ساختار مکان شناسانه خالص گراف زمانی بر می‌گردد که تابع وزن برای همه  $(i, j) \in E$  برابر  $w_{ij} = 1$  باشد و ماتریس مجاورت را در این مورد نیز پوشش دهد. توجه

1 Recursive bisection  
2 Electrical Connections  
3 Topological  
4 Weighted Vertex Degree

5 Pure connectivity of the network  
6 Admittance  
7 Line resistance  
8 Reactance  
9 Electrical distance  
10 Average power flow  
11 Real  
12 Internal structure  
13 Island  
14 Real symmetric matrix

اگر وزن‌ها با راکتانس‌های معکوس خط برابر باشند (با استفاده از مدل شبکه DC)، لاپلاسین نرمالیزه نشده با ماتریس ادمیتانس گره (با صرف نظر از سوسپتانس‌های شنت<sup>۲</sup>) برابر می‌شود. وقتی وزن‌ها با توان عبوری برابر باشند نمی‌توان از لاپلاسین تعبیری در حوزه برق داشت.  
۲. لاپلاسین نرمالیزه شده: لاپلاسین نرمالیزه شده ورودی‌های  $d_i$  نا صفر دارد. یعنی:

$$[L_n]_{ij} = \begin{cases} 1 & \text{if } i=j \\ \frac{-w_{ij}}{\sqrt{d_i}\sqrt{d_j}} & \text{if } i \neq j \text{ and } (i,j) \in E \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

لاپلاسین نرمالیزه شده مستقل از مقیاس<sup>۳</sup> می‌باشد و برای دسته‌بندی نیز مفید تر است (بخش ۴-الف را ببینید).

### ج. مقادیر ویژه لاپلاسین

ماتریس مقادیر ویژه لاپلاسین (نرمالیزه شده یا نرمالیزه نشده) باید ویژگی‌های اساسی زیر را داشته باشد [۱۴].

الف. تمام مقادیر ویژه اعداد حقیقی نا منفی<sup>۴</sup> هستند.

ب. صفر یک مقدار ویژه است و با تعداد اجزای متصل در گراف (جزایر) تعیین می‌شود.

مقادیر ویژه  $L$  را به صورت  $0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$  و مقادیر ویژه  $L_n$  را نیز به صورت  $0 = v_1 \leq v_2 \leq \dots \leq v_n$  می‌نویسیم. از ویژگی بی‌دانیم که  $v_p > 0$  یا  $\lambda_p > 0$ ، اگر و تنها اگر  $G$  متصل باشد. مقادیر ویژه  $L$  به مقیاس وابسته هستند و حد بالا ندارند (همه وزن‌ها و مقادیر ویژه در یک عدد اسکالر ضرب می‌شوند). با این حال مقادیر ویژه  $L_n$  در نامساوی  $0 \leq v_i \leq 2$  برای همه  $i$ ها صدق می‌کنند.

### ۳- خوشه بندی طیفی

الف. معرفی:

با دسته‌بندی به دنبال شناسایی تعدادی از گره‌ها در گراف (به نام دسته) هستیم به طوری که گره‌ها در یک دسته اتصالات قوی (با در نظر گرفتن وزن یال‌ها) و در دسته‌های دیگر اتصالات ضعیفی دارند. با دسته‌بندی طیفی به دنبال نوعی دسته‌بندی هستیم که از مقادیر و بردارهای ویژه لاپلاسین استفاده کند. استدلال نظری برای دسته‌بندی طیفی در بخش (۳-ج) ارائه می‌شود. ایده کلی ابتدا استفاده از  $k$  بردار ویژه لاپلاسین برای دادن مختصات هندسی<sup>۵</sup> به گره‌های  $N$  در  $R^k$  برای  $2 \leq k \ll N$  می‌باشد. یعنی این مختصات  $N$  ردیف ماتریس  $N \times k$  است که ستون‌هایش  $k$  بردار ویژه از کوچکترین مقادیر ویژه<sup>۶</sup> هستند. سپس داده‌های حاصل با استفاده از برخی الگوریتم‌های دسته‌بندی استاندارد توسعه یافته و برای نقاط در فضای اقلیدسی<sup>۷</sup> دسته‌بندی می‌شوند. بعداً می‌گوییم دسته‌بندی طیفی می‌کوشد تا

$K$  دسته متوازن<sup>۸</sup> را تقریباً با حجم<sup>۹</sup> برابر ایجاد نماید. برای سنجش کیفیت تقسیم بندی‌ها به کمیتی<sup>۱۰</sup> به نام گسترش زیر گراف نیاز داریم.  
ب. گسترش زیر گراف:

زیر گراف، گرافی است که چند گره به صورت  $S \subset V$  دارد، گره‌ها  $i \in S$  و یال‌ها  $(i,j) \in E$  هستند به طوری که  $i,j \in S$ . بر اساس تعاریف موجود در مهندسی برق، جزیره زیر گراف  $G$  با زیر مجموعه‌ای  $(S)$  از گره‌ها ( $V$ ) خواهد بود. (توجه کنید جزایر قبلی نباید متصل شوند). ضمناً فقط جزایر مهم را در نظر بگیریم ( $S \neq \emptyset, S \neq V$ ). بر اساس تعریف، حجم جزیره احاطه شده توسط مجموعه  $S \subset V$ ، مجموع مقادیر گره‌هایش می‌باشد:

$$vol(S) = \sum_{i \in S} d_i \quad (3)$$

توجه کنید حجم  $S$  با وزن‌های مربوط به گره‌ها و نه فقط اعداد (تعداد) آنها در  $S$  مشخص شده است. حدود که جمع وزن‌های اتصالات بین گره‌های موجود در  $S$  و گره‌هایی که در  $S$  نیستند (یعنی جمع وزن خطوط متصل کننده  $S$  به بقیه شبکه) می‌باشد.

$$\partial(S) = \sum_{i \in S, j \notin S} w_{ij} \quad (4)$$

برای سنجش کیفیت جزیره از نسبت اندازه حدود<sup>۱۱</sup>  $\partial(S)$  به حجم  $S$  استفاده خواهد شد. به این دلیل گسترش  $S$  به صورت زیر تعریف شد:

$$\phi(S) = \frac{\partial(S)}{vol(S)} \quad (5)$$

هرچه گسترش  $S$  کم تر باشد سبب می‌شود تا جزیره بهتر بتواند استحکام اتصالات<sup>۱۲</sup> بین گره‌هایش ( $vol(S)$  بزرگ) و اتصالات ضعیف به سایر نقاط شبکه ( $\square(S)$  کوچک) را بهبود بخشد. به عنوان نمونه برای مفاهیم حدود، حجم و گسترش، شکل (۲) و جدول (۲) را ببینید. توجه کنید به حداقل رساندن حدود  $\partial(S)$  به جای کاهش نسبت، پاسخ‌های نا کافی<sup>۱۳</sup> را ارائه می‌کند به طوری که جزیره با کم ترین حدود در شبکه فقط دارای گرهی با کوچکترین درجه است (شرح شکل (۲) را ببینید). از این رو به حداقل رساندن حدود بدون محدودیت‌های اضافی<sup>۱۴</sup> سبب افزایش پاسخ‌های کاذب<sup>۱۵</sup> می‌شود.

ج. نامساوی cheeger :

نامساوی Cheeger یکی از استدلال‌های نظری اصلی روش دسته‌بندی طیفی است. چون کیفیت جزیره  $S \subset V$  را با میزان گسترش  $\phi(S)$  می‌سنجیم، کیفیت  $k$  بخشی<sup>۱۶</sup> (یعنی یک بخش در  $k$  دسته)  $U S_1 \cup \dots \cup U S_k = V$  را با بیشترین گسترش در جزیره‌ها اندازه گیری خواهیم کرد.

$$\max_{1 \leq i \leq k} \phi(S_i) \quad (6)$$

(با توجه به گسترش<sup>۱۷</sup>) بهترین  $k$  بخشی در کل گراف، حداقل  $k$  بخش موجود در گراف است:

$$\rho_G(k) = \min_{\emptyset \neq S_1, \dots, S_k \subseteq V} \max_{1 \leq i \leq k} \phi(S_i) \quad (7)$$

<sup>8</sup> K balanced clusters

<sup>9</sup> Volume

<sup>10</sup> Quantity

<sup>11</sup> Ratio of the size of the boundary

<sup>12</sup> Highly interconnected

<sup>13</sup> Inadequate solutions

<sup>14</sup> Additional constraints

<sup>15</sup> Spurious solutions

<sup>16</sup> K-partition

<sup>17</sup> Expansion

<sup>1</sup> Nodal admittance matrix

<sup>2</sup> Shunt susceptances

<sup>3</sup> Scale-independent

<sup>4</sup> Nonnegative real

<sup>5</sup> Geometric coordinates

<sup>6</sup> Smallest K eigenvalues

<sup>7</sup> Euclidean space

برای ابر نقاط<sup>۹</sup> موجود در فضای اقلیدسی، برای شناسایی دسته‌بندی‌ها با  $k$  جایگیری طیفی که مقادیر آنها تغییر کرده است می‌باشد و اگر  $k > 3$  باشد تجسم دشوار است. انتخاب استاندارد، انتخاب الگوریتم  $k$ -mean است. با این حال این روند از دیدگاه کنترل کلی فرآیند دسته‌بندی چند عیب دارد. اولاً باید از قبل تصمیم بگیرید می‌خواهید چند دسته داشته باشید، معمولاً برای داشتن  $k$  دسته از جایگیری طیفی در  $R^k$  استفاده می‌شود. دوماً این روش اجازه کد گذاری<sup>۱۰</sup> ساختار دسته‌بندی کل شبکه اصلی را نمی‌دهد، ولی در عوض می‌توان خروجی یک بخشی شبکه را کد گذاری کرد. سوماً الگوریتم دسته‌بندی در فضای اقلیدسی، این مسئله که نقاط نشان دهنده گره‌ها در گراف هستند را نادیده می‌گیرد، یعنی اطلاعات یال‌ها را در نظر نمی‌گیرد.

برای حل این معایب، به جای الگوریتم استاندارد  $k$ -mean [۱۷] از دسته‌بندی سلسله مراتبی<sup>۱۱</sup> استفاده می‌کنیم. این روش ساختار شبکه را در میزان وضوح مختلف و به صورت همزمان نشان می‌دهد که آنها با استفاده از دیاگرام درختی<sup>۱۲</sup> کد گذاری شدند.  $dendrogram$  یک دیاگرام درختی (همانند شکل (۶)) است که ساختار سلسله مراتبی را در جایگذاری طیفی کد گذاری می‌کند. برگ‌های پایینی مبین گره‌های منفرد (شین‌ها) است و به عنوان دسته‌های اولیه با اندازه‌ی ۱ در نظر گرفته شده‌اند. در هر درخت، نزدیک‌ترین دسته‌ها با هم ادغام می‌شوند تا فاصله بین دسته‌ها به کم‌ترین فاصله بین نقاط جفت موجود در دسته برسد. یک دیاگرام درختی نمایان‌گر ساختار سلسله مراتبی شبکه در همه سطوح و در واقع همان  $k$  بخشی، در یک لحظه است، برای چند  $k$  می‌توان با برش دیاگرام درختی در سطح  $k$  ام از ریشه،  $k$  بخشی را بازیابی کرد (برای مثال شکل (۶) را ببینید).

برای برطرف کردن عیب سوم، بر اساس جایگیری طیفی به گراف یک معیار دادیم که اطلاعات یال‌ها را در نظر می‌گیرد. فرض کنید  $u_i \in R^k$  مختصات نرمالیزه شده مربوط به  $i$  امین گره باشند. اگر گره‌های  $i$  و  $z$  در گراف متصل باشند، سپس  $||u_i - u_z|| = dist(i, z)$  برابر فاصله یال قرار گرفته می‌شوند، به طور کلی  $dist(i, j)$  حداقل فاصله در مسیرهای متصل به  $i$  و  $z$  در گراف است، این فاصله در مسیر مجموع فواصل یال‌ها است. مزیت اصلی این است که فقط اجازه ادغام جزیره‌هایی وجود دارد که حداقل با یک یال بهم وصل شده‌اند. به این "معیار گذاری در گراف با جایگیری"<sup>۱۳</sup> گفته می‌شود. آخرین اصلاحی که باید انجام شود این است که باید فاصله را به جای خط مستقیم روی یک کره محاسبه کرد، چون ابر نقطه را روی کره  $k-1$  بعدی دسته‌بندی می‌شود. در ادامه در الگوریتم پیشنهادی به شرح و توجیه گام چهارم پرداخته می‌شود. بدین ترتیب یک دسته‌بندی با وابستگی به مقیاس بدست آورده شد که اجازه می‌دهد تا به داخل دسته‌ها رفته تا ساختار ریز تر و دقیق تر آنها نشان داده شود. این سطح جدید به صورت کامل کنترل ساختار کاری صحیح شبکه را در وضوح متفاوت نشان می‌دهد. به طور خاص خروجی الگوریتم ما فقط یک بخش از شبکه نیست (اگر چه می‌توان چنین بخش‌هایی را فوراً بدست آورد) بلکه در عوض به صورت یک ساختار دسته‌بندی سلسله مراتبی با

این کمیت، ثابت گسترش  $k$  راهه<sup>۱</sup> گراف  $G$  نام دارد. روش جست و جو فراگیر<sup>۲</sup> برای یافتن بخش بهینه (مانند فوق) برای شبکه‌های بزرگ از نظر محاسباتی عملی نیست و در واقع مشکلات جست و جو موجود در گراف این چینی عموماً سخت هستند [۱۵]. دسته‌بندی طیفی یک راه حل تقریبی<sup>۳</sup> ارائه می‌کند که پیچیدگی محاسباتی در آن فقط به محاسبه  $k$  مقدار ویژه و بردار ویژه ماتریس لاپلاسیان بستگی دارد (برای ماتریس متقارن و حقیقی با اندازه‌ی  $N$ ). نامساوی Cheeger می‌سنجد این راه حل تقریبی چقدر به راه حل بهینه<sup>۴</sup> نزدیک است و این پاسخ بهینه چقدر کلیت دارد، یعنی این گراف چقدر برای  $k$  بخشی مناسب است. برای  $k=2$  نامساوی Cheeger برابر است با [۱۴]:

$$\frac{v_z}{z} \leq \rho_G(\tau) \leq \sqrt{2} v_z \quad (8)$$

و این نامساوی کلاسیک اخیراً برای  $k \geq 2$  در مجانب زیر تعمیم یافته است:

$$\frac{v_k}{z} \leq \rho_G(k) \leq O(k^{\frac{1}{2}}) \sqrt{v_k} \quad (9)$$

در واقع سمت راست معادله‌های (۸-۳) و (۹-۳) با استفاده از خوشه‌بندی طیفی و با در نظر گرفتن لاپلاسیان نرمالیزه شده بدست آمده‌اند. سمت چپ آنها نیز مقدار  $\frac{v_k}{z}$  است که نشان دهنده حداقل مقدار گسترش هر  $k$  بخشی شبکه می‌باشد. بنابراین هرچه  $v_k$  کوچکتر باشد پاسخ تقریبی دسته‌بندی طیفی به پاسخ بهینه نزدیک تر و جامع تر خواهد بود. اگر  $v_k$  بزرگ باشد گراف دارای  $k$  بخشی مناسب نخواهد بود. مهم این است دسته‌بندی یک شبکه در چند دسته و به حداقل رساندن گسترش کلی (۳-۶) معادل با اقدامات لازم برای توازن حجم همه دسته‌ها باشد.

#### ۴- استدلال<sup>۵</sup> روش دسته‌بندی طیفی سلسله مراتبی

الف. استدلال:

اساس روش دسته‌بندی طیفی استفاده از  $k$  بردارهای ویژه اولیه لاپلاسیان گراف برای دادن مختصات به گره‌ها در فضای اقلیدسی<sup>۶</sup>  $R^k$  بعدی است. به این  $k$  جایگیری طیفی<sup>۷</sup> می‌گویند و نشان دهنده دسته‌بندی در  $k$  جزیره است. گونه‌های مختلفی از این روش کلی وجود دارند. اولین انتخاب پیش رو این است که آیا از لاپلاس نرمالیزه شده استفاده شود یا نرمالیزه نشده. استفاده از لاپلاس نرمالیزه شده  $L_n$  به ویژه برای شبکه‌هایی با مقادیر دلخواه چند دلیل دارد [۱۶] و در شبیه سازی متوجه می‌شویم که بردارهای ویژه لاپلاس نرمالیزه شده از نظر دسته‌بندی پاسخ‌های بهتری دارند. این گام اول و دوم در الگوریتم مان را توجیه می‌کند (بخش ۴-ب را ببینید). مشخص است هنگام استفاده از بردارهای ویژه  $L_n$  یک گام نرمالیزاسیون اضافی پس از جایگیری طیفی<sup>۸</sup> وجود دارد که در آن بردارهای ویژه باید نرمالیزه شوند تا طول آنها یک شود. این گام سوم در بخش ۴-ب را توجیه می‌کند. انتخاب دوم، الگوریتم دسته‌بندی

<sup>1</sup> K-way expansion constant

<sup>2</sup> Brute-force approach

<sup>3</sup> Approximate solution

<sup>4</sup> Optimal solution

<sup>5</sup> Methodology

<sup>6</sup> Euclidean space

<sup>7</sup> Spectral k-embedding

<sup>8</sup> Spectral embedding

<sup>9</sup> Point clouds

<sup>10</sup> Encode

<sup>11</sup> Hierarchical clustering

<sup>12</sup> Dendrogram

<sup>13</sup> The metric induced on the graph by the embedding

جایگیری طیفی است که با استفاده از دیاگرام درختی کد گذاری و تجسم شده است.

هنوز می توان بعد  $k$  که در آن شبکه جایگذاری می شود را انتخاب کرد که به نوبه خود به تعداد دلخواه دسته‌هایی وابسته است که درصدد یافتن آنها هستیم. خود این عدد می تواند یک مسئله مستقل باشد، برای نمونه حجم دلخواه دسته‌ها، هم چنین باید بعد بهینه را برای جایگیری طیفی بدست‌آید. یک معیار رایج استفاده از فواصل ویژه<sup>۱</sup> است، یعنی اختلاف بین دو مقدار ویژه متوالی (به یاد بیاورید که اگر  $v_k$  کوچک باشد یک  $k$  بخشی مناسب وجود خواهد داشت). چون به دنبال مقادیر ویژه کوچکتر هستیم، از نسبت اختلاف مقدار ویژه به اندازه اش استفاده می‌شود:

$$\gamma_k = \frac{v_{k+1} - v_k}{v_k} \quad (k \geq 2) \quad (10)$$

مقدار  $\gamma_k$  بالا (برای مثال شکل (۳) را ببینید) سبب می‌شود تا شبکه حداقل به  $k$  جزیره تقسیم شود و این با جایگیری طیفی در  $k$  بعد نشان داده خواهد شد. توجه کنید دسته‌بندی سلسله مراتبی ساختار زیر دسته را در هر دسته، فراتر از یک  $k$  بخشی حفظ می‌کند. بنابراین اگر هیچ معیاری وجود نداشته باشد از مقدار بالای  $\gamma_k$  استفاده خواهد شد تا برای جایگیری طیفی (گام صفرم) بعد انتخاب شود.

ب. الگوریتم:

$G = (V, E, W)$  اگر گراف دارای وزن بدون جهت با  $N$  گره باشد.

گام صفر (بعد طیف):

$k \ll N$  را با مقدار بالای  $\gamma_k$  فاصله ویژه نسبی در رابطه (۱۰-۳) انتخاب کنید.

گام اول (لاپلاسین نرمالیزه شده):

اگر  $W = (w_{ij})$  ماتریس وزن یال‌ها باشد، تعریف کنید  $D = \text{diag}(d_i)$  که  $d_i = \sum_j w_{ij}$

$$L_n = D^{-1/2} L D^{-1/2} \quad \text{و} \quad L = D - W$$

گام دوم (جایگیری طیفی):

از  $k$  بردار ویژه لاپلاسین نرمالیزه شده  $L_n$  برای دادن مختصات به گره‌های گراف در  $\mathbb{R}^k: x_1, \dots, x_N$  استفاده کنید. (اگر  $X$  ماتریس  $N \times k$  باشد که ستون‌هایش  $k$  بردارهای ویژه اولیه باشند، می‌توان گفت  $x_i$  امین ردیف  $X$  است).

گام سوم (نرمالیزاسیون):

بردارها نرمالیزه کنید تا طول آنها در  $\mathbb{R}^k$  برابر ۱ شود:

$$u_i = \frac{x_i}{\|x_i\|}, \quad 1 \leq i \leq N \quad (11)$$

(این منتهی به طرح شعاع<sup>۲</sup> روی کره  $K-1$  بصورت  $S^{k-1} = \{u \in \mathbb{R}^k \text{ such that } \|u\| = 1\}$  می‌شود).

گام چهارم (دسته‌بندی سلسله مراتبی):

دیاگرام درختی نقاط  $\mathbb{R}^k \subset S^{k-1} \subset \{u_1, \dots, u_N\}$  با توجه به معیار داده شده به گراف با جایگیری که در بخش (۴-الف) توضیح داده شده است را رسم کنید.

ج. پیچیدگی الگوریتم:

بیشترین زمان محاسباتی مربوط به گام دوم، محاسبه مقادیر ویژه ماتریس با اندازه  $N$  است که معکوب می‌باشد. از MATLAB استفاده شد در آن الگوریتم  $\text{eigs}(L, k, \text{diff})$  را برای محاسبه مقادیر ویژه ماتریس متقارن و متناهی و مثبت به کار گرفته شد. در عمل این محاسبه در موارد آزمایشی کسری از ثانیه زمان می‌برد که معمولاً به خاطر کوچک بودن ماتریس‌ها است (برای دیدن زمان‌های متوسط بیش از ۱۰۰ آزمایش به جدول (۱) مراجعه کنید).

جدول (۱) زمان محاسبه و مکان شناسی مقادیر ویژه

متوسط زمانی	تعداد شاخه‌ها	تعداد شین‌ها	
شبه ۰/۰۰۰۵ < ثانیه	۴۶	۳۹	شبکه ۳۹ شینه IEEE
شبه ۰/۰۰۴ < ثانیه	۱۷۹	۱۱۸	شبکه ۱۱۸ شینه IEEE
شبه ۰/۳۳ ثانیه	۱۰۴۷	۸۱۵	(ادمیتانس) شبکه GB
شبه ۰/۳۷ ثانیه	۱۰۲۱	۷۹۰	توان عبوری شبکه GB

د. تعبیر فیزیکی فاصله‌ی طیف:

$k$  جایگیری طیفی تابع فاصله بین شین‌ها را به ما می‌دهد و مقایسه آن با امپدانس معادل تون<sup>۴</sup> که معیاری استاندارد برای فاصله الکتریکی است، جالب می‌باشد. بعداً این را می‌توان با معکوس کردن ماتریس ادمیتانس محاسبه کرد (یعنی لاپلاسین با مقادیر برابر با ادمیتانس‌ها) که نشان دهنده اتصال با فاصله طیفی<sup>۵</sup> در زمانی است که ادمیتانس‌ها به عنوان وزن‌ها مورد استفاده قرار گیرند. به طور کلی فاصله بدست آمده با معکوس کردن لاپلاسین اغلب با جایگیری طیفی نرمالیزه نشده همبسته است ولی رابطه ریاضی مشخص و مستقیمی وجود ندارد. در عمل ما ضرایب همبستگی خطی بین فاصله الکتریکی و طیفی را برای ادمیتانس نرمالیزه نشده بر اساس لاپلاسین پیدا می‌کنیم: ۰.۸۹ و ۰.۸۲ و ۰.۸۷ به ترتیب برای سیستم‌های ۳۹ شینه IEEE و ۱۱۸ شینه IEEE و موارد آزمایشی GB برای انتخاب‌های مناسب  $k$  می‌باشند. انتظار می‌رود اتصالات قوی بین شین‌ها سبب فواصل الکتریکی کوچکتر شود. تفاوت عمده این است که فاصله طیفی گره‌ها را طوری سازماندهی می‌کند که  $k$  دسته نشان داده شوند؛ از این رو با توجه به پارامتر دسته‌بندی  $k$  یک فاصله جهانی است. از نظر ریاضی امپدانس تونن تمام بردارهای ویژه لاپلاسین نرمالیزه نشده (محاسبه تمام مقادیر و بردارهای ویژه ماتریس قطری با محاسبه معکوسش برابر است) را در نظر می‌گیرد در حالی که جایگیری طیفی فقط از  $k$  کوچکترین مقادیر ویژه استفاده می‌کند. توجه کنید که فاصله طیفی به جایگیری طیفی در فضای اقلیدسی مربوط می‌شود که به ما دید فضایی یا هندسی می‌دهد در حالی که امپدانس تونن یک معیار عددی خالص و بدون اطلاعات اضافی در مورد فضا است.

#### ۵- نتایج شبیه سازی کامپیوتری و تحلیل آن‌ها

الف. سیستم آزمایشی ۳۹ شینه IEEE:

روش پیشنهادی در سیستم آزمایشی ۳۹ شینه IEEE با جزئیات کامل بیان می‌شود. در بخش بعد روش پیشنهادی را در سیستم‌های دیگر به کار

<sup>4</sup> Thévenin's equivalent impedance

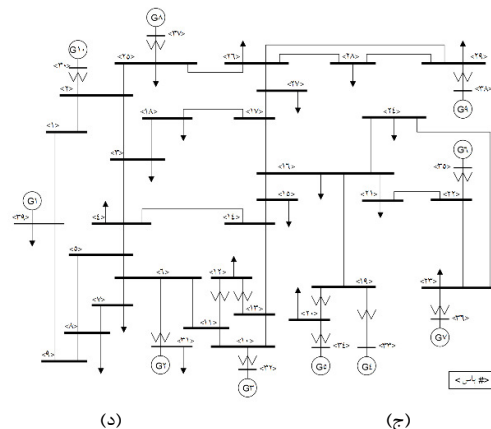
<sup>5</sup> Spectral distance

<sup>1</sup> Eigengaps

<sup>2</sup> Normalization

<sup>3</sup> Radial projection

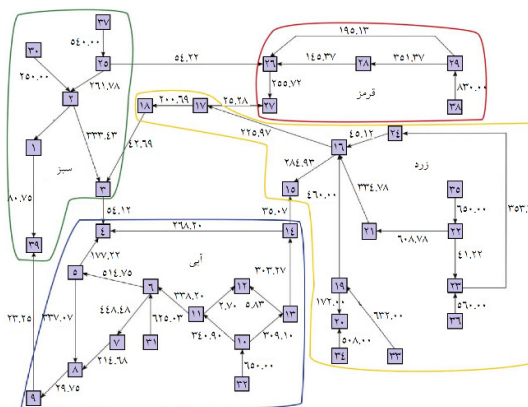
گرفته می‌شود و یافته‌های اصلی نیز بیان می‌گردند. از دسته‌بندی طیفی مبتنی بر لاپلاسین توان عبوری کار شروع می‌شود و به مقایسه نتایج با ادمیتانس مبتنی بر لاپلاسین پرداخته می‌شود. تمام نتایج عددی روی متلب نسخه R2015a (۶۴ بیتی) با یک پردازنده ۲.۵ گیگاهرتز با رم ۴ گیگابایتی اجرا شدند.



شکل ۱ نمایش شبکه ۳۹ شینه IEEE

ب. لاپلاسین توان عبوری:

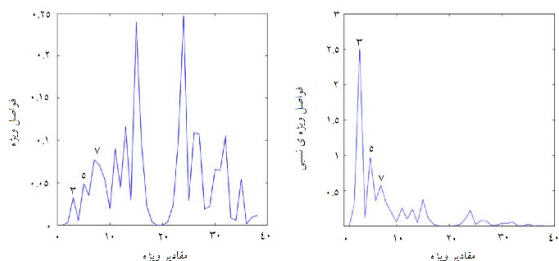
تحلیل دسته‌بندی طیفی سیستم‌های آزمایشی ۳۹ شینه IEEE با توان عبوری DC در شکل (۲) نشان داده شده است. ابتدا روی ابعاد ۲ و ۳ تمرکز می‌کنیم که در این ابعاد می‌توانیم جایگیری‌ها را تجسم کنیم. به طور کلی از فواصل ویژه نسبی (رابطه (۱۰-۳)) برای تعیین بعد طیفی بهینه  $k$  استفاده می‌کنیم و اگر  $k > 3$  نمایش دیاگرام درختی برای نشان دادن ساختار داخلی سلسله مراتبی سخت تر می‌شود.



شکل ۲ توان عبوری در یک مورد ۳۹ شینه در IEEE و دسته‌بندی به ۴ ناحیه به روش پیشنهادی

همانطور که در شکل (۲) مشاهده می‌کنید جزایری با رنگ‌های سبز، قرمز، زرد و آبی مشخص شده است. جدول (۲) را برای حدود، حجم‌ها و نسبت گسترش این جزایر مشاهده نمایید. توجه کنید مطابق شکل دسته‌بندی طیفی جهت توان عبوری را نادیده می‌گیرد. مشاهده می‌کنید جزیره‌ای که شین ۱۲ در آن قرار دارد کم‌ترین حدود  $\Phi(S) = 8/53$  اما بیشترین میزان گسترش  $\Phi(S) = 100\%$  را دارد و حدود آن در تمام این حجم قرار دارند.

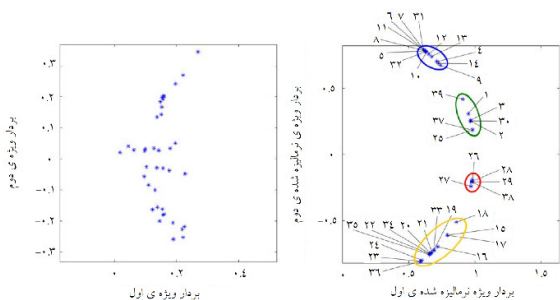
گام صفر (بعد طیفی). بعد طیفی  $k$  را از نقاط پیک در فواصل ویژه نسبی (شکل (۳)) انتخاب می‌شود. در این مورد پیک در  $k = 3.5$  قرار دارد که نشان دهنده استفاده ابعاد ۳ یا ۵ برای نشان دادن ساختار دسته‌بندی اصلی است.  $k=3$  قرار داده می‌شود، اما برای ابعاد ۲ و ۳ برای رسیدن به اهداف مورد نظر تحلیل انجام می‌شود



شکل ۳ فواصل ویژه (چپ) و فواصل ویژه نسبی (راست) برای توان عبوری مبتنی بر لاپلاسین در مورد ۳۹ شینه IEEE

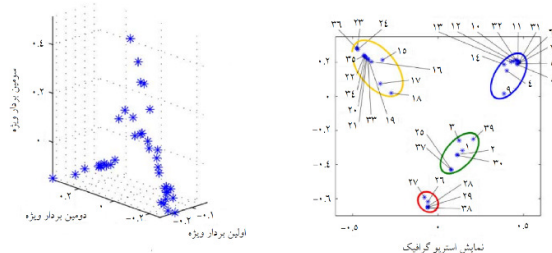
گام اول و دوم (لاپلاسین نرمالیزه شده، جایگیری طیفی). لاپلاسین نرمالیزه شده را با توجه توان عبوری ایجاد می‌شود، بردار ویژه محاسبه می‌گردد و از بردارهای ویژه  $k$  برای قرار دادن مختصات در شین‌ها در  $\mathbb{R}^k$  برای  $k=2,3$  بهره برده می‌شود. این شبکه را در فضای  $\mathbb{R}^2$  (شکل (۴)) و در فضای  $\mathbb{R}^3$  (شکل (۵)) چپ نشان می‌دهد. در حال حاضر می‌توان پدیدار شدن دسته‌بندی شعاعی، یعنی دسته‌بندی نقاط موجود در راستای نیم خط که نسبت به مبدا پدیدار می‌شوند را مشاهده کرد.

گام سوم (نرمالیزاسیون). مختصات نرمالیزه می‌شوند تا تمام بردارها برابر با یک شوند، یعنی با یک دایره واحد ( $k=2$ ) یا کره واحد ( $k=3$ ) کار می‌شود. این روند نقاط را در راستای خطوط شعاعی از مبدا قرار می‌دهد و مبنی ساختار دسته‌بندی شعاعی فوق است. در شکل (۴) جایگیری طیفی شعاعی را قبل (چپ) و بعد (راست) از نرمالیزه کردن مشاهده می‌کنید. با بررسی تصاویر می‌توان پدیدار شدن ساختار دسته‌بندی سلسله مراتبی با چهار دسته و چند زیر دسته مربوط به نقاط داده را مشاهده نمود (شکل (۴) راست). برای  $k=3$  نقاط در کره دو بعدی بدست آورده شد، برای تجسم بهتر روی صفحه استروگرافیک انجام می‌شود (شکل (۵) راست).



شکل ۴ جایگیری طیفی با ابعاد ۲ برای نمونه آزمایشی ۳۹ شینه IEEE و توان عبوری

همانطور که در شکل (۴) مشاهده می کنید نمودار سمت راست، جایگیری نرمالیزه شده (نقاط روی دایره قرار دارند) با نقاطی که با شماره شین نام گذاری شدند را نشان می دهد و ۴ دسته نیز مشخص شده اند (از بالا به پایین: آبی، سبز، قرمز و زرد).



شکل ۵ جایگیری طیفی با ابعاد ۳ برای نمونه آزمایشی ۳۹ شین *IEEE*، توان عبوری (چپ) و نمایش استریوگرافیک (راست).

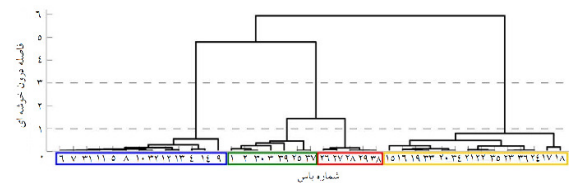
همانطور که در شکل ۵ نمایان است به صورت بصری می توان ۴ دسته (با رنگ های آبی، سبز، قرمز و زرد به ترتیب به صورت ساعتگرد از گوشه بالا سمت چپ) را مشاهده کرد که مشابه جایگیری طیفی با ابعاد ۲ (شکل (۴) راست) هستند. عنوان ها شماره شین هستند. نمودار سمت راست با نرمالیزه کردن بردارها برای داشتن طول واحد و نمایش استریوگرافیک از نقطه (۱۰۰۰) بدست آمده است.

دسته بندی کوچکتر را می توان مشاهده کرد (بخش (۵-ب) را ببینید). توجه کنید بردارهای ویژه نرمالیزه شده نقاطی روی یک نیم دایره (شکل (۴) راست) و روی یک نیم کره (نشان داده نشده است) دارند و این به خاطر بردار ویژه  $\pm \sqrt{d_1}, \dots, \sqrt{d_n}$  است از این رو مختصات اولیه همه نقاط نیز همیشه علامت یکسانی دارند.

گام چهارم (دسته بندی سلسله مراتبی) ابعاد انتخاب شده تا حد کافی کوچک هستند تا امکان دسته بندی با بررسی بصری (شکل (۴) راست و (۵-۴) راست) فراهم باشد، دیاگرام درختی اطلاعات دسته بندی مشابهی فراهم می کند (شکل (۶-۵) را ببینید).

با مراجعه به شبکه اصلی بررسی می شود که این ۴ جزیره به معنی باشند (شکل (۲))، در واقع متصل بوده و حدود کوچکی برای حجم هایشان (جدول (۲)) داشته باشند.

مقایسه نسبت های گسترش محاسبه شده با حدود نظری پایین تر  $V_k/2$  برای  $k=3,4$  (بخش (۳-ج) را ببینید) جالب است: به ترتیب ۰.۰۶۵ و ۰.۰۲۲۶. یعنی  $k$  بخشی بهینه (که برای زمان چند جمله ای<sup>۱</sup> قالب کلی ندارد) گسترشی حداقل برابر با ۰.۶۵٪ و ۲.۲۶٪ بترتیب برای  $k=3,4$  دارد و با مقادیر محاسبه شده (در چند میلی ثانیه) یعنی ۲.۰۵٪ و ۵.۰۳٪ به ترتیب، مغایرت دارد.



شکل ۶ دیاگرام جایگیری طیفی با بعد ۳ برای سیستم آزمایشی ۳۹ شین *IEEE* با مقادیر توان عبوری

در شکل (۶) ارتفاع جزیره های آبی، سبز، زرد و قرمز (چپ به راست روی محور افقی) در ۱ بازیابی شده است در حالی که برش در ارتفاع ۳ سبب ادغام جزیره های سبز و قرمز برای ۳ بخشی مناسب می شود (شکل (۲) و جدول (۲)). خطوط عمودی بلند در دیاگرام درختی نشان دهنده دسته بندی مناسب در آن سطح می باشند.

جدول (۲) تقسیم کردن به ۴ جزیره

شین ها (اندازه)	
۳۱-۳۲، ۴-۱۴	آبی
۳۰، ۳۱، ۳۷	سبز
۲۶-۲۹، ۳۸	قرمز
۱۵-۲۴، ۳۳-۳۶	زرد
۳-۱، ۲۵-۳۰، ۳۷	سبز + قرمز

#### ب. نشان دادن اتصالات داخلی شبکه

از جایگیری طیفی می توان استفاده بیشتری برد: در شکل های (۴) راست) و (۵) راست) با بررسی می توان ساختار دسته های فرعی غنی را مشاهده کرد. برای مثال مکان نسبی شین های منفرد نشان دهنده این است که آیا به مرکز جزیره تعلق دارند یا به اطراف آن (برای مثال شین ۹ در شکل (۵) راست را ببینید). با بررسی دیاگرام درختی (شکل (۶)) مشاهده می شود ادغام جزایر سبز و قرمز سه بخشی مناسبی ارائه می کند (شکل (۱) و جدول (۲)).

عدم پوشش ساختار دسته بندی های فرعی هدف اصلی انتخاب دسته بندی سلسله مراتبی در روش پیشنهادی است. ساختار سلسله مراتبی کامل شبکه در دیاگرام درختی کد گذاری می شود. اگرچه دیاگرام درختی در داده ها مقیاس های مختلف را می گیرد اما انتخاب های مختلف بعد جایگیری طیفی با توجه به این که ابعاد جایگیری طیفی مختلف بر اساس  $k$  بخشی متوازن پیکر بندی های متفاوتی دارند ( $k$  دسته تقریباً با اندازه  $vol(V)/k$ ) بخش های نسبتاً متفاوتی را نشان خواهد داد.

#### ج. لاپلاسیان ادمیتانس

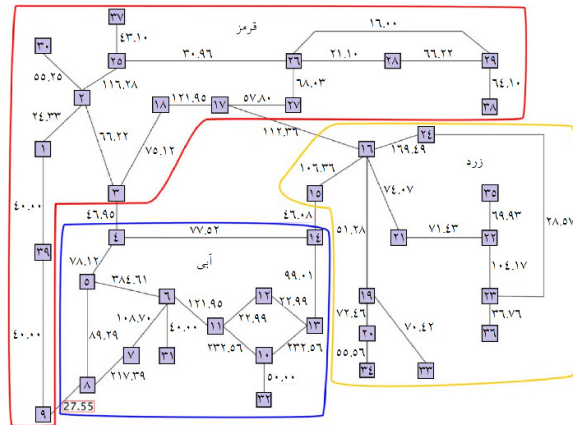
برای لاپلاسیان مبتنی بر ادمیتانس تحلیل مشابهی انجام می شود. ورودی آن مشابه بوده اما مقادیر متفاوت هستند، ادمیتانس های خط نیز در شکل (۷) نشان داده شده اند. تقسیم بندی شبکه با استفاده از جایگیری طیفی با بعد ۳ نشان داده می شود. با توجه به این که ادمیتانس های خط و توان عبوری همبسته هستند (با ضریب همبستگی خطی ۰.۶۶) سه جزیره حاصل مشابه جزیره مربوط به لاپلاس توان عبوری در شکل (۲) هستند، در واقع محدود به توپولوژی شبکه می باشند (ادمیتانس و توان عبوری بین دو شین صفر است، اگر به صورت فیزیکی به هم متصل نشده باشند). از طرف دیگر با توجه به این که مقادیر توان عبوری کوچکتر هستند نسبت گسترش بیشتر از مقادیر مربوط به توان عبوری است (۹.۳۴٪ قرمز، ۸.۰۰٪ زرد، ۳.۲۸٪ آبی) (انحراف استاندارد<sup>۲</sup> ۲۱۲.۵ در مقایسه با ۶۸.۶ برای مقادیر (وزن های

<sup>۲</sup> Periphery

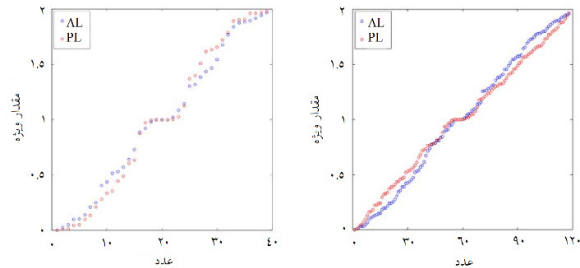
<sup>۳</sup> Standard deviation

<sup>۱</sup> Polynomial time

ادمیتانس). این را می‌توان در مقادیر ویژه لاپلاس (شکل ۸) چپ مشاهده کرد.



شکل ۷ ادمیتانس‌های خط در سیستم آزمایشی ۳۹ شینه IEEE و سه دسته‌بندی (با رنگ‌های زرد، قرمز و آبی) که به روش پیشنهادی مشخص شده‌اند.



شکل ۸ مقادیر ویژه برای توان عبوری مبتنی بر (PL) و ادمیتانس مبتنی بر (AL) لاپلاسین نرمالیزه شده سیستم آزمایشی ۳۹ شینه IEEE (چپ)، سیستم آزمایشی ۱۱۸ شینه IEEE (راست).

همانطور که در شکل ۸ نمایان است منحنی  $S$  شکل از ویژگی‌های مقادیر ویژه  $0 \leq v_k \leq 2$  و  $\sum_k v_k = N$  و فراوانی  $v = 1$  (تعداد بالای مقادیر ویژه اغلب به خاطر تقارن شبکه اصلی می‌باشد [۱۸]) حاصل شده است. هرچه این شکل  $S$  بارز تر باشد، تقسیم بندی برای  $k$ های مور انتظار در شبکه کوچکتر خواهد بود. از این رو نمودارها برای روش توان عبوری در مقایسه با روش ادمیتانس تقسیم بندی‌های بهتری در سیستم آزمایشی ۳۹ شینه پیشبینی می‌کنند و این در سیستم آزمایشی ۱۱۸ شینه بالعکس است.

در واقع کران پایین نظری  $v_k/2$  برای لاپلاسین مبتنی بر ادمیتانس بالا تر هستند:  $۱.۳۴\%، ۱.۶۸\% و ۲.۵۰\%$  بترتیب برای  $k=۲،۳،۴$ . تحلیل دقیق تر دیاگرام درختی به ما اجازه کشف ساختار دسته‌بندی دیگر را می‌دهد، برای مثال بزرگترین جزیره (آبی) را می‌توان به دو بخش ۱-۳، ۹، ۱۸، ۳۰، ۳۹ و ۴-۱۰، ۱۴-۳۱ تقسیم کرد.

#### ۶- نتیجه گیری

یک الگوریتم کارآمد پیشنهاد شد که از دسته‌بندی طیفی و سلسله مراتبی برای نشان دادن ساختار اتصالات داخلی یک شبکه قدرت و با در نظر گرفتن انتخاب پارامتر الکتریکی مربوط به خطوط انتقال استفاده می‌کند. برای نشان دادن روش پیشنهادی در سه سیستم کوچک، متوسط و بزرگ از توان عبوری

و ادمیتانس‌های خط استفاده کردیم. به علاوه الگوریتم مان را سنجیدیم و از آن به عنوان ابزار تقسیم بندی<sup>۱</sup> و یک روش رقابتی<sup>۲</sup> استفاده کردیم. استفاده از توان عبوری به عنوان مقادیر (وزن‌ها) در کاربردهایی که مقادیر پویا هستند مفید تر است و هم چنین به شرایط کاری واقعی وابسته می‌باشند. یک کاربرد می‌تواند جزیره‌ای کردن پیشگیرانه باشد، دسته‌بندی می‌کوشد تا قطع توان عبوری<sup>۳</sup> ناشی از برش در اتصالات بین دسته‌ها را به حداقل برساند، می‌توان از آن در اتاق کنترل به عنوان روش تجسم<sup>۴</sup> شرایط کاری واقعی استفاده کرد. گروه‌های مختلف در سیستم بسته به بار گذاری شان به هم نزدیک تر می‌شوند یا از هم فاصله می‌گیرند، از این رو طرح و ساختار بار گذاری توان عبوری را نشان می‌دهند. وقتی مطلوب است اتصالات شبکه با فاصله الکتریکی بیان شوند دسته‌بندی طیفی سلسله مراتبی را می‌توان با استفاده از ادمیتانس‌های خط انجام داد. بنابراین این دسته‌بندی ساختار داخلی ثابت شبکه را نشان می‌دهد. به ویژه در ابعاد بزرگتر از ۳ از دیاگرام‌های درختی برای سازماندهی<sup>۵</sup> و نمایش بصری ساختار سلسله مراتبی شبکه استفاده گردید. این فراتر از الگوریتم  $k$ -mean استاندارد است و روش پیشنهادی ساختار سلسله مراتبی کاری شبکه را فراتر از یک تقسیم بندی ساده نشان می‌دهد. دیاگرام درختی اطلاعات مربوط به بخش‌های فرعی<sup>۶</sup> (یا ادغام شده) موجود در جزیره‌های کوچکتر (بزرگتر) را حفظ می‌کند.

باور داریم که جایگیری طیفی و دیاگرام درختی مربوطه، شیوه صحیح نمایش اتصالات داخلی شبکه با توجه به مجموعه‌ی مشخصی از مقادیر (وزن‌های) یال‌ها می‌باشند. با کاهش فضای جست و جو و ارائه دیدگاه هندسی<sup>۷</sup> در مورد سلسله مراتب کاری<sup>۸</sup> شبکه می‌توان این را سرآغازی برای تحلیل بیشتر به حساب آورد. با کاوش دیاگرام درختی و شبکه اصلی می‌توان بر اساس کاربردی که در نظر است جزئیات را به خوبی پیدا کرد. نظریه ریاضی اساسی را با هدف بیان پتانسیل بالای روش دسته‌بندی طیفی در هنگام استفاده در حوزه<sup>۹</sup> مهندسی قدرت به طور مفصل توضیح داده شد.

#### ۷- نوآوری

یک نرم افزار تحت وب بر اساس روش دسته‌بندی طیفی سلسله مراتبی طراحی شد. این نرم افزار فعلا قادر به انجام روش پیشنهادی بر روی شبکه‌ی ۳۹ شینه IEEE می‌باشد. این نرم افزار از کمپایلر آنلاین متلب استفاده می‌کند. عملکرد این نرم افزار بدین گونه است که ابتدا مقادیر و وزن‌های توان عبوری (یا ادمیتانس و یا وزن خالص شبکه) را از کاربر دریافت می‌کند و سپس بر اساس آن ماتریس لاپلاسین نرمالیزه را تشکیل می‌دهد. بر اساس فواصل ویژه بدست آمده از مقادیر ویژه نرمالیزه لاپلاسین، بعد جایگیری طیفی  $k$  را بدست می‌آورد و سپس بر اساس بعد جایگیری طیفی، تعداد مقادیر ویژه کوچک و بردارهای ویژه مربوط به آن را بدست می‌آورد. بر اساس بردارهای ویژه جایگیری طیفی را روی دایره (۲ بعدی) و کره (سه بعدی) نشان می‌دهد. در نهایت دیاگرام درختی روش دسته‌بندی سلسله مراتبی را

1 Partitioning tool  
2 Competing technique  
3 Disruption of power flows  
4 Visualizing  
5 Organize  
6 Subdivisions  
7 Geometric insight  
8 The functional hierarchy  
9 Context



- [15] D. S. Johnson and M. R. Garey, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. San Fransisco, CA, USA: Freeman, 1979.
- [16] U. VonLuxburg, M. Belkin, and O. Bousquet, "Consistency of spectral clustering," *Ann. Statist.*, pp. 586-555, 2008.
- [17] A. Y. Ng, M. I. Jordan, Y. Weiss, "On spectral clustering: Analysis and an algorithm," *Adv. Neural Inf. Process. Syst.*, vol. 2, pp. 856-849, 2002.
- [18] B. D. MacArthur and R. J. Sánchez-García, "Spectral characteristics of network redundancy," *Phys. Rev. E*, vol. 80, no. 2, p. 026117, 2009.

نمایش می‌دهد. این نرم افزار کدهای مربوطه را پردازش می‌کند و نمایش آن‌ها به عهده‌ی کمپایلر<sup>۱</sup> آنلاین متلب می‌باشد. برای طراحی این نرم افزار از زبان اسکریپت نویسی *PHP* و *HTML* و *CSS* استفاده کردیم.

#### ۷- پیشنهادات

با توجه به قابلیت‌های روش دسته‌بندی طیفی سلسله مراتبی در سیستم‌های قدرت و دسته‌بندی شبکه‌ی انتقال به چند ناحیه و جزیره‌ی جدا، می‌توان از آن در اتاق فرمان سازمان برق جهت مشاهده‌ی جزایری با مصرف بالای برق استفاده کرد. بدین صورت که مشخصه‌های توان عبوری در شبکه‌ی مورد نظر را، نرم افزار طراحی شده در قسمت نو آوری، از طریق بانک‌های اطلاعاتی و یا کاربر شبکه‌ی هوشمند دریافت می‌کند و دسته‌بندی طیفی را را انجام می‌دهد. از این جهت پیشنهاد می‌شود که از این روش جهت مانیتورینگ شبکه در اتاق فرمان استفاده شود. پیشنهاد دیگر ما این است که جهت بررسی اتصال خالص شبکه و شین‌ها نیز از این روش استفاده شود. می‌توان این روش را گسترش داد و مقادیر مختلف توان عبوری را با این روش سنجید و ذخیره کرد و سپس برای این روش یک چگالی احتمال تعریف کرد. که در این صورت این روش قادر به پیش بینی توان عبوری و دسته‌بندی آن در ساعات مختلف مصرف می‌شود. پیشنهاد دیگر ما این است که این روش را برای شبکه‌های توزیع استفاده کرد و شبکه‌های توزیع نسبتاً گسترده را به چند جزیره دسته‌بندی کرد و سپس با قرار دادن بریکرها در محل‌های مناسب باعث جلوگیری از خاموشی شد. می‌توان این روش را برای کنترل و مانیتورینگ شبکه برق در پالایشگاه‌ها، پتروشیمی‌ها و... استفاده کرد.

#### مراجع

- [1] G. Kron, *Diakoptics: The Piecewise Solution of Large-Scale Systems*. New York, NY, USA: MacDonald, 1963, vol. 2.
- [2] H. H. Happ, *Piecewise Methods and Applications to Power Systems*. New York, NY, USA: Wiley, 1980.
- [3] J. H. Chow, *Time-Scale Modeling of Dynamic Networks With Applications to Power Systems*. New York, NY, USA: Springer-Verlag, 1982.
- [4] F. Wu, "Solution of Large-scale networks by tearing," *IEEE Trans. Circuits Syst.*, vol. 23, no. 12, pp. 713-706, Dec. 1976.
- [5] S. Blumsack, P. Hines, M. Patel, C. Barrows, and E. Cotilla Sanches, "Defining power network zones from measures of electrical distance," in *Proc. 2009 IEEE Power & Energy Society General Meeting*, 2009, pp. 1-8.
- [6] S. McArthur, "The autonomic power system-network operation and control beyond smart grids," in *Proc. 2012 IEEE Innovative Smart Grid Technologies (ISGT Europe)*, 2012, pp. 7-1.
- [7] G. Xu and V. Vittal, "Slow coherency based cutset determination algorithm for large power systems," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 877-884, May 2010.
- [8] K. Sun, D. Z. Zhang, and Q. Lu, "Splitting strategies for islanding operation of large-scale power systems using OBDD-based methods," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 923-912, May 2003.
- [9] C. Wang, B. Zhang, Z. Hao, J. Shu, P. Li, and Z. Bo, "A novel real-time searching method for power system splitting boundary," *IEEE Trans. Power Syst.*, vol. 25, no. 4, pp. 1909-1902, Nov. 2010.
- [10] H. Li, G. W. Rosenwald, J. Jung, and C. C. Liu, "Strategic power infrastructure defense," *Proc. IEEE*, vol. 93, no. 5, pp. 933-918, May 2005.
- [11] L. Ding, F. M. Gonzalez-Longatt, P. Wall, and V. Terzija, "Two-step spectral clustering controlled islanding algorithm," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 84-75, Feb. 2013.
- [12] L. Liu, X. Chen, D. Luo, Y. Lu, G. Xu, M. Liu, "HSC: a spectral clustering algorithm combined with hierarchical method," *Neural Network World* vol. 6, pp. 499-521, 2014.
- [13] J. R. Lee, S. Oveis Gharan, and L. Trevisan, "Multi-way spectral partitioning and high-order Cheeger inequalities," in *Proc. 44<sup>th</sup> ACM Symp. Theory of Computing*, 2012, pp. 1130-1117.
- [14] F. R. Chung, *Spectral Graph Theory*, ser. CBMS Regional Conference Series in Mathematics. Providence, RI, USA: American Mathematical Society, 1997, no. 92.

<sup>1</sup> Compiler

