دانشگاه صنعتی شریف
دانشکده مهندسی کامپیوتر

مهندسی کامپیوتر

پروژه آمار و احتمال مهندسی

نگارش

پرهام رضایی ۴۰۰۱۰۸۵۴۷

امیرحسین عزیزی ۴۰۰۱۰۵۱۲۲

محمدعرفان سلیما ۴۰۰۱۰۵۰۱۴

استاد درس

دکتر شریفی زارچی

بهمن ۱۴۰۱

فهرست مطالب

فصل ۱

بخش های پروژه

۱-۱ ز هرچه رنگ تعلّق پذیرد آزاد است!

- تئوری ۱:
همه رئوس یک خوشه، بجز درایه مربوط به خود راس با یکدیگر سطر و یا ستون یکسانی دارند. که همه رئوس یک خوشه چون به هم یال دارند درایه برابر ۱ دارند. درواقع می‌توان نوعی افراز برای سطر ها و ستون ها بدست آورد. که در هریک همه درایه ها بجز راس با خودش یک اند و این زیرماتریس ها با یکدیگر اشترک ندارند. درایه های باقی مانده همگی صفر می‌باشند.

- تئوری ۲: دیگر نمیتوان افراز کرد چراکه امکان وجود اشتراک بین خوشه های تعیین شده است. اما میتوان تعدادی زیرماتریس استخراج کرد که هر یک خواص ذکر شده در تئوری قبلی را دارند یعنی درایه های برابر بجز راس به خود دارند ولی این زیرماتریس ها این بار میتوانند با هم اشتراک داشته باشند.

- تئوری۳: برای دو عضو که اشتراک ژانرهای مشتریکشان ژانرهای $i_1, i_2, \ldots, i_n$ باشد احتمال هم سلیقه بودن آن ها برابر با متمم هم سلیقه نبودنشان است پس:

$$P = 1 - \Pi_{i=1}^{n}(1 - p_i)$$

با عضو بودن در جوامع خاص زیاد اشتراک جوامع آنها زیاد شده و ضریب های $1 - p_i$ ی بیشتری ضرب شده که موجب کوچک شدن عبارت منفی و در نتیجه بزرگ شدن احتمال یال داشتن می شود.

هرچه اشتراک دو نفر بیشتر باشد این بدان معناست که سلیقه آن دو به هم نزدیک تر است. پس احتمال یال داشتن منطقا باید بیشتر شود.

در این نوع مدل سازی صرفا به اشتراک آنها توجه شده و به گروه های عدم اشتراکشان دقت نکردیم با این که اگر دو نفر فقط در یک گروه مشترک باشند و اینکه این دو نفر فقط در همین خوشه عضو باشند در عمل نشان از سلیقه مشترک کمتری دارد اما در این نوع مدل احتمال یکسانی برای هم سلیقگی دارند.

- تئوری ۴:

$$< A_۱, A_۲ >= \sum F_{۱,i} F۲, i$$

$$\implies e^{-<A_۱, A_۲>} = e^{-\sum F_{۱,i} F_{۲,i}}$$

$$\implies P_{i,j} = ۱ - e^{-<A_i, A_j>}$$

پس با بالا رفتن اشتراک تعداد یم در جایگاه های برابر ستون ها بیشتر میشود پس ضرب داخلی آنها بیشتر شده که به موجب آن طرف راست تساوی افزایش یافته پس احتمال ارتباط که طرف چپ تساوی است نیز افزایش می‌یابد.

- تئوری ۵:

$$P(A_{i,j} = ۱ \mid F) = P_{i,j},\ P(A_{i,j} = ۰ \mid F) = ۱ - P_{i,j}$$

$$\implies P(A_{i,j} \mid F) = A_{i,j} P_{i,j} + (۱ - A_{i,j})(۱ - P_{i,j})$$

$$\implies P(A \mid F) = \Pi_{۱ \leqslant i < j \leqslant n} P(A_{i,j} \mid F)$$

$$\implies \updownarrow(F) = \sum_{۱ \leqslant i < j \leqslant n} \log(A_{i,j} P_{i,j} + (۱ - A_{i,j})(۱ - P_{i,j}))$$

- تئوری ۶:

چراکه با حرکت در راستای گرادیان قدر مطلق افزایش می‌یابد. پس شهودا منطقی است مشابه رگرسیون خطی هر مرحله سعی کنیم گرادیان را کم کنیم تا به قله های محلی برسیم.

- تئوری ۷:

$$\frac{dl(F)}{dF_{u,i}} = \sum_k \frac{F_{u,i} F_{k,i} (۱ - ۲A_{u,k}) e^{-<F_u, F_k>}}{e^{-<F_u, F_k>}(۱ - ۲A_{u,k}) + A_{u,k}}$$

- تئوری ۸: حال برای مقدار به جای برنولی از نمایی استفاده می کنیم.

$$A_{i,j} = x \rightarrow P =< F_i, F_j > e^{-<F_i,F_j>x}$$

$$\frac{dl}{dF_{u,i}} = \sum_k \frac{F_{u,i}F_{k,i}e^{-<F_u,F_k>A_{u,k}} + -F_{u,i}F_{k,i} < F_u, F_k > e^{-<F_u,F_k>A_{u,k}}}{< F_u, F_k > e^{-<F_u,F_k>A_{u,k}}}$$

۱-۲ گناه بخت پریشان و دست کوته ماست!

- تئوری ۹:

$$Q = \begin{bmatrix} ۰/۶ & ۰/۱ & \dots & ۰/۱ \\ ۰/۱ & ۰/۶ & \dots & ۰/۱ \\ \vdots & \vdots & \ddots & \vdots \\ ۰/۱ & ۰/۱ & \dots & ۰/۶ \end{bmatrix}$$

$$A_{i,j} = \begin{cases} ۱ & Q[Z[i], Z[j]] \\ ۰ & ۱ - Q[Z[i], Z[j]] \end{cases}$$

- تئوری ۱۰:

خیر. درایه های قطر اصلی که فرد با خودش است حتما همسلیقگی وجود دارد که یعنی ۱ ولی در ساخت ما ۰.۶ است. همچنین درایه های متقارن باید با هم برابر باشند اما در ساخت ما مستقل حساب شده اند که به دلیل ذکر شده بسیار نادرست است.

- تئوری ۱۱: طبق فرض ما بجز قطر اصلی و مثلث پایینی که با مثلث بالایی برابر است . مابقی مستقل از یکدیگرند که پس تعداد آنها:

$$\binom{n}{۲}$$

است.

- تئوری ۱۲:

$$P(A \mid Z) = \Pi_{j<i}(۱ - A[i,j])(۱ - Q[Z[i], Z[j]]) + A[i,j]Q[Z[i], Z[j]]$$

- تئوری ۱۳:

$$\sum_{j>i} \log((۱ - A[i,j])(۱ - Q[Z[i], Z[j]]) + A[i,j]Q[Z[i], Z[j]])$$

۱-۳ او نه خیال است و نه طیف!

- تئوری ۱۴:

برای یک درایه P_{ij} برابر است با p اگر طرفدار ژانر موافق باشند و q در غیر این صورت. حال داریم که

$$A_{i,j} = \begin{cases} ۱ & P_{i,j} \\ ۰ & ۱ - P_{i,j} \end{cases} \quad (۱-۱)$$

درواقع یعنی متغیر تصادفی برنولی می‌باشد.

- تئوری ۱۵:

لیست ژانر مورد علاقه افراز را لیست مرتب Z می نامیم.

$$\mathbb{E}[A_{i,j}] = ۱ \times P_{i,j} + ۰ \times (۱ - P_{i,j}) = P_{i,j}$$

بنابر این داریم که

$$W_{i,j} = \begin{cases} p & Z[i] = Z[j] \\ q & Z[i] \neq Z[j] \end{cases} \quad (۱-۲)$$

- تئوری ۱۶:

بدون از دست دادن کلیت با جابجایی اعضا داریم که:

$$W = \begin{vmatrix} p & p & q & q \\ p & p & q & q \\ q & q & p & p \\ q & q & p & p \end{vmatrix} \quad (۱-۳)$$

می‌توان با قراردهی معادله صفر شدن دترمینان بردار ویژه های زیر و مقدار ویژه های نظیرشان را مشاهده کرد.

$$(۱, ۱, ۱, ۱) \to p + q$$

$$(۱, ۱, -۱, -۱) \to p - q$$

$$(۰, ۰, ۱, -۱) \to ۰$$

$$(۱, -۱, ۰, ۰) \to ۰$$

- تئوری ۱۷:

در حالت کلی بدون از دست دادن کلیت با جابجایی اعضا و قراردهی در نیمه اول و دوم داریم که (دقت کنید مقدار ویژه ها ثابتند):

تعریف می کنیم

$$q_{a\times b}, p_{a\times b} \in M_{a\times b}(\mathbb{C})$$

که تمام درایه های آن q باشد. مشابها برای p .

پس داریم که چون طبق گفته کانال تعداد برابری دارند، به شکل زیر است:

$$W = \begin{vmatrix} p_{k\times k} & q_{k\times n-k} \\ q_{n-k\times k} & p_{n-k\times n-k} \end{vmatrix} \qquad (۴-۱)$$

حال کلا دو ستون متمایز داریم که در حالت کلی ضرییبی از هم نیستند. پس رنک این ماتریس دو است. بدین سبب تعداد مقدار ویژه های ناصفر آن نیز باید دو باشد. حال می‌توان مشاهده کرد که دو بردار زیر بردار ویژه اند با مقدار ویزه نظیر

$$p+q \to \left|1_{n\times 1}\right| \qquad (۵-۱)$$

$$p-q \to \begin{vmatrix} 1_{k\times 1} \\ -1_{n-k\times 1} \end{vmatrix} \qquad (۶-۱)$$

حال برای فضای نال این ماتریس داریم که بردارهایی که دو عضو متوالی یک و منفی یک و مابقی صفرند به شرطی که عضو k ام نباشد. ماتریس را صفر میکند. این بردار ها مستقل اند. زیرا برای صفر شدن جمعشان باید ضریب هر دو بردار متوالی در تایk اول و n-k تای دوم برای ایجاد جمع صفر برای یک و منفی یک برابر باشد. و ضریب اخرین عضو k تای اول و اولین عضو n-k تای دوم برای صفر شدن (چون درایه یکتا اند) باید صفر باشند. پس همگی باید صفر باشند که معادل استقلال خطی است.

- تئوری ۱۸:

$$\begin{bmatrix} k(p+q) & 0 & \dots & 0 \\ 0 & k(p+q) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & k(p+q) \end{bmatrix} \qquad (۷-۱)$$

- تئوری ۱۹ :

$$\begin{bmatrix} (k-1)p+q & -p & \dots & -q & \dots & -q \\ -p & (k-1)p+q & \dots & -q & \dots & -q \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ -q & -q & \dots & -p & \dots & (k-1)p+q \end{bmatrix} \qquad (۸-۱)$$

حال برای مقادیر و بردار های ویژه داریم که:

$$L_w v = \lambda v \implies D_w v - W v = \lambda v \implies W v = (k(p+q) - \lambda)v$$

پس بردار ویژه های W نیز هستند. برعکس این معادلات نیز درست است. پس بردار ویژه های یکسان ولی با مقدار ویژه های متفاوت دارند. پس

$$۰ \to k(p+q) - \lambda = ۰ \implies \lambda = k(p+q)$$

$$k(p+q) \to k(p+q) - \lambda = k(p+q) \implies \lambda = ۰$$

$$k(p-q) \to k(p+q) - \lambda = k(p-q) \implies \lambda = ۲kq$$

پس $n-۲$ تا مقدار ویژه $k(p+q)$ ، یکی ۰ و یکی $۲kq$ دارد. که بردار ویژه های معادل همان بردار ویژه های W اند. یعنی برای $k(p+q)$ ها فقط دو عضو ناصفر متوالی یک و منفی یک. برای ۰ همگی یک و برای $۲kq$ نیمه نخست یک و نیمه دوم منفی یک.

- تئوری ۲۰:

$$u_۱ = (۱, ۱, \dots, ۱)$$

$$u_۲ = (۱, \dots, ۱, -۱, \dots, -۱)$$

پس نیمه نخست نقاط به صورت (۱،۱) و نیمه دوم بصورت (۱،-۱) می باشند. با قرار دارن ترشولد برابر صفر برای مولفه دوم میتوان خوشه بندی کرد که معادل خوشه بندی مدنظر است.

- تئوری ۲۱: برای خوشه بندی دو بخشی روی بردار ویژه دوم ترشولد ۰ قرار میدهیم. پس خطا زمانی است که u در بین اکسپکتد و اصلی تفاوت علامت داشته باشد. نرمال اگر در نظر بگیریم. هر درایه اکسپکتد

$$\pm \frac{۱}{\sqrt{n}}$$

می‌باشد. بدین سبب هر بار که علامت تغییر کند به جمع ارائه شده در سوال حداقل $\frac{۱}{n}$ افزوده می‌شود. پس کران بالای تعداد آن‌ها

$$\frac{k}{n} \leqslant \frac{C}{\mu^۲} \implies .k \leqslant \frac{Cn}{\mu^۲}$$

اگر قرار دهیم $n = t\mu^۲$ داریم که به احتمال $۱ - ۴e^{-t\mu^۲}$ حداکثر Ct خطا دارد. و برای اینکه احتمال اپسیلون شود. $۱ - ۴e^n = ۱ - \epsilon \implies n = ln(\frac{\epsilon}{۴})$

- تئوری ۲۲:

مشابه الگوریتم ارائه شده ماتریس لاپلاسین را ساخته و سپس بردار ویژه های آنها را بدست آورده و برحسب مقدار ویژه ها صعودی سورت می کنیم. خواص لاپلاسین و مقدار ویژه ها در عکس های زیر گزارش شده اند. سپس برای خوشه بندی روی k خوشه، الگوریتم kmean را روی k بردار ویژه در سورت انجام شده، لیبل های نظیر را بدست می‌آوریم.

$$L = D - W.$$

An overview over many of its properties can be found in Mohar (1991, 1997). The following proposition summarizes the most important facts needed for spectral clustering.

Proposition 1 (Properties of L) *The matrix L satisfies the following properties:*

1. *For every vector $f \in \mathbb{R}^n$ we have*

$$f'Lf = \frac{1}{2}\sum_{i,j=1}^{n} w_{ij}(f_i - f_j)^2.$$

2. *L is symmetric and positive semi-definite.*
3. *The smallest eigenvalue of L is 0, the corresponding eigenvector is the constant one vector $\mathbb{1}$.*
4. *L has n non-negative, real-valued eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$.*

Proof.
Part (1): By the definition of d_i,

$$f'Lf = f'Df - f'Wf = \sum_{i=1}^{n} d_i f_i^2 - \sum_{i,j=1}^{n} f_i f_j w_{ij}$$
$$= \frac{1}{2}\left(\sum_{i=1}^{n} d_i f_i^2 - 2\sum_{i,j=1}^{n} f_i f_j w_{ij} + \sum_{j=1}^{n} d_j f_j^2\right) = \frac{1}{2}\sum_{i,j=1}^{n} w_{ij}(f_i - f_j)^2.$$

Part (2): The symmetry of L follows directly from the symmetry of W and D. The positive semi-definiteness is a direct consequence of Part (1), which shows that $f'Lf \geq 0$ for all $f \in \mathbb{R}^n$.
Part (3): Obvious.
Part (4) is a direct consequence of Parts (1) - (3). □

Note that the unnormalized graph Laplacian does not depend on the diagonal elements of the adjacency matrix W. Each adjacency matrix which coincides with W on all off-diagonal positions leads to the same unnormalized graph Laplacian L. In particular, self-edges in a graph do not change the corresponding graph Laplacian.

The unnormalized graph Laplacian and its eigenvalues and eigenvectors can be used to describe many properties of graphs, see Mohar (1991, 1997). One example which will be important for spectral clustering is the following:

Proposition 2 (Number of connected components and the spectrum of L) *Let G be an undirected graph with non-negative weights. Then the multiplicity k of the eigenvalue 0 of L equals the number of connected components $A_1, \ldots, A_k$ in the graph. The eigenspace of eigenvalue 0 is spanned by the indicator vectors $\mathbb{1}_{A_1}, \ldots, \mathbb{1}_{A_k}$ of those components.*

Proof. We start with the case $k = 1$, that is the graph is connected. Assume that f is an eigenvector with eigenvalue 0. Then we know that

$$0 = f'Lf = \sum_{i,j=1}^{n} w_{ij}(f_i - f_j)^2.$$

Unnormalized spectral clustering

Input: Similarity matrix $S \in \mathbb{R}^{n \times n}$, number k of clusters to construct.
- Construct a similarity graph by one of the ways described in Section 2. Let W be its weighted adjacency matrix.
- Compute the unnormalized Laplacian L.
- **Compute the first k eigenvectors $u_1, \ldots, u_k$ of L.**
- Let $U \in \mathbb{R}^{n \times k}$ be the matrix containing the vectors $u_1, \ldots, u_k$ as columns.
- For $i = 1, \ldots, n$, let $y_i \in \mathbb{R}^k$ be the vector corresponding to the i-th row of U.
- Cluster the points $(y_i)_{i=1,\ldots,n}$ in $\mathbb{R}^k$ with the k-means algorithm into clusters $C_1, \ldots, C_k$.

Output: Clusters $A_1, \ldots, A_k$ with $A_i = \{j | \, y_j \in C_i\}$.

Normalized spectral clustering according to Shi and Malik (2000)

Input: Similarity matrix $S \in \mathbb{R}^{n \times n}$, number k of clusters to construct.
- Construct a similarity graph by one of the ways described in Section 2. Let W be its weighted adjacency matrix.
- Compute the unnormalized Laplacian L.
- **Compute the first k generalized eigenvectors $u_1, \ldots, u_k$ of the generalized eigenproblem $Lu = \lambda Du$.**
- Let $U \in \mathbb{R}^{n \times k}$ be the matrix containing the vectors $u_1, \ldots, u_k$ as columns.
- For $i = 1, \ldots, n$, let $y_i \in \mathbb{R}^k$ be the vector corresponding to the i-th row of U.
- Cluster the points $(y_i)_{i=1,\ldots,n}$ in $\mathbb{R}^k$ with the k-means algorithm into clusters $C_1, \ldots, C_k$.

Output: Clusters $A_1, \ldots, A_k$ with $A_i = \{j | \, y_j \in C_i\}$.

۱-۴ من از دیار حبیبم نه از بلاد غریب!

- تئوری ۲۳: $p^m(۱-p)^{\binom{n}{۲}-m}$
- تئوری ۲۴: $\frac{۱}{\binom{\binom{n}{۲}}{m}}$
- تئوری ۲۵: $\sum_{k=۰}^{\min(m,\frac{S}{۵})} \binom{m}{k}p^k(۱-p)^{m-k}\binom{S-m}{\frac{S}{۵}-k}(۱-p)^{\frac{S}{۵}-k}p^{S-m-\frac{S}{۵}+k}$
- تئوری ۲۶:

$\mathbb{E} = \binom{n}{2}p$

- تئوری ۲۷: احتمال هرکدام از یالهای متصل به آن را حمع میکنیم درنتیجه

$$(n-۱)p$$

- تئوری ۲۸:

$$۱ - \sum_{i=۰}^{(n-۱)p-۱} (۱-p)^{n-i}p^i$$

- تئوری ۲۹:

امید ریاضی تعداد روابط هم سلیقگی دارای خاصیت تراگذری برابر میشود با:

$\binom{n}{3}\mathbb{E}[I_1] = \binom{n}{3}p^3$

امید ریاضی تعداد روابط هم سلیقگی دارای خاصیت زنجیره ای برابر میشود با:

$\binom{n}{3}\mathbb{E}[I_1] = \binom{n}{3}3p^2(1-p)$

- تئوری ۳۰:

$\mathbb{P} = \frac{p^3}{p^3+3p^2(1-p)}$

- تئوری ۳۱: برابر تعداد مثلث هایی است که راس درون آنهاست. اگر برای کل جمع بزنیم طبق خطی بودن امید ریاضی داریم که:

$n\mathbb{E}[I_1] = \mathbb{E}[I_1 + \cdots + I_n] = 3\binom{n}{3}p^3 \implies E[I_1] = \binom{n-1}{2}p^3$

- تئوری ۳۲:

$\mathbb{P}[I_{u,v} = 1] = (1-p^2)^{n-2}$

- تئوری ۳۳:

$\mathbb{E}[X_n] = \binom{n}{2}(1-p^2)^{n-2}$

- تئوری ۳۴:

$\mathbb{P}[X_n \geqslant 1] \leqslant \mathbb{E}[X_n] = \binom{n}{2}(1-p^2)^{n-2}$

$\lim_{n\to\infty} \frac{n(n-1)(1-p^2)^{n-2}}{2} = \lim_{n\to\infty} \frac{(2n-1)k^2}{2k^n \ln k} = \lim_{n\to\infty} \frac{2k^2}{2k^n \ln k^2} = 0, \frac{1}{1-p2} = k$

- تئوری ۳۵: در نتیجه وقتی که n عدد بزرگی باشد ،هر دو راس یک دوست مشترک دارند پس قطر بزرگ برابر ۲ میشود.
  قطر گراف برای nهای بزرگ به p وابسته نیست.
  نتیجه تئوری با نتیجه شبیه سازی تطابق دارد.

۱-۵ سل المصانع رکبا تهیم فی الفلوات

- تئوری ۳۶:

$$P_{i,j} = \begin{cases} \frac{۱}{d_i} & j \in v(i) \\ ۰ & j \notin v(i) \end{cases} \qquad (۹-۱)$$

- تئوری ۳۷:

$$A = DP \implies P = D^{-1}A$$

- تئوری ۳۸:

$$P^{۲}_{i,j} = \sum_{k=۰}^{n} P_{i,k} Pk, j$$

که برابر احتمال رسیدن از i به j است . که روی تمام مسیر های ممکن برحسب احتمال آنها جمع کرده ایم.

- تئوری ۳۹: p^t

- تئوری ۴۰:

برای جفت i ، j مسیرهای بین این دو را از دو سمت دوگانه شماری میکنیم. در هر مسیر وزن هایی که یک تقسیم بر درجه اند در یکدیگر ضرب می شوند. در مسیر های از .i به j داریم که وزن یالهای میانی همگی ضرب شده و یک تقسیم بر درجه i نیز ضرب می شود. برای همین مسیر از j به i داریم که وزن یالهای میانی ضرب شده و صرفا اولین ضریب معادل i برای j می شود. پس نسبت این دو برابر $\frac{d_i}{d_j}$ است. و با سیگما بستن روی همه مسیر های بین این دو داریم که

$$\frac{P_{i,j}}{P_{j,i}} = \frac{\frac{۱}{d_i}}{\frac{۱}{d_j}} = \frac{d_j}{d_i}$$

- تئوری ۴۱: برای P_{ij} برای T های کوچک احتمال کمتری نسبت به T های بزرگتر دارد زیرا در شبکه هم سلیقگی افراد با مسیر کوتاه‌تری به هم وصل هستند ولی برای P_{kj} برای T های کوچک مقدار بسیار کوچکی دارد زیرا احتمال اینکه دو نفر در یک شبکه بزرگ هم سلیقه باشند بسیار کم است و با افزایش T مقدار بزرگ‌تری پیدا میکند و ماکسیمم می‌شود و با

افزایش مجدد T کاهش می‌یابد زیرا در یک گراف طول مسیر حداکثر n است و مسیر طولانی با طول مشخص احتمال کمی برای شکل گرفتن دارد.

- تئوری ۴۲:

$$\mathcal{P}^t_{C,k} = \frac{(\sum_{i \in C} P^t_{i,k})}{(\mid C \mid)}$$

فرم کلی

$$r_{C_۱,C_۲} = \sqrt{\frac{(\mathcal{P}_{C_۱,k} - \mathcal{P}_{C_۲,k})^۲}{d(k)}}$$

البته میتوان به فرم

$$r_{C_۱,C_۲} = \sqrt{\frac{(\mathcal{P}_{C_۱,C_k} - \mathcal{P}_{C_۲,C_k})^۲}{d(C_k)}}$$

که احتماله ها بصورت زیر تعریف شوند نیز نمایش داد

$$P_{C_i,C_j} = \frac{۱}{\mid C_i \mid\mid C_j \mid} \sum_{a \in C_i, b \in C_j} P^t_{a,b}$$

- تئوری ۴۳: باید یالهای داخل خوشه از تعداد یال های آن به رئوس بیرون بیشتر باشد پس از تابع Q زیر استفاده میکنیم.

$$Q = \sum_{k=۱}^{n} \left(\frac{l_k}{m} - \left(\frac{d_k}{۲m} \right)^۲ \right)$$

که m_k تعداد یالهای درون خوشه و d_k جمع درجات رئوس خوشه و m برابر تعداد یال های گراف در مجموع می‌باشد.