



دانشکده علوم ریاضی
گروه ریاضی کاربردی

پروژه کارشناسی ریاضیات و کاربردها گرایش گراف و درجه بندی

عنوان

رتبه بندی و درجه بندی راس های گراف

پژوهشگر

یگانه امینی

استاد راهنما

دکتر مجید گازر

تابستان 1403

چکیده:

یکی از مسایل مهم در نظریه گراف ، علوم کامپیوتر و شبکه های اجتماعی، مشخص کردن اهمیت راس های یک گراف است. یکی از روش ها، رتبه بندی است که بر پایه گام برداری بنا شده است. هدف ما در این مقاله ، توضیح الگوریتم رتبه بندی به دو شکل متمرکز و توزیع شده است. بدین منظور نخست مفهوم رتبه بندی در الگوریتم محاسبه ان را به صورت متمرکز توضیح میدهیم.

واژگان کلیدی: روش مونت کارلو، گام برداری تصادفی ، معیار های مرکزیت، سیستم های توزیع شده.

فهرست مطالب

فصل 1: مقدمه

۱.۰.۱ مروری بر فصل های دیگر 6

فصل 2: تعاریف

۱.۲

۱.۱.۲ گراف 7

۱.۱.۲ مثال 8

۲.۱.۲ مثال 8

۲.۱.۲ طوقه 10

۳.۱.۲ گراف متناهی 10

۴.۱.۲ گراف ساده 10

۵.۱.۲ گراف مرکب 10

۶.۱.۲ درجه راس ها 10

۱.۱.۲ قضیه 10

فصل 3: رتبه بندی راس های گراف

۱.۳ سرآغاز و انگیزه

۱.۱.۳ مثال 11.....

۲.۱.۳ مثال 12.....

۱.۱.۳ تعریف 12.....

۲.۳ مفاهیم و نتایج مقدماتی

۱.۲.۳ مرکزیت 12.....

۱.۲.۳ تعریف 13.....

۱.۲.۳ مثال 13.....

۲.۲.۳ زنجیر مارکوف و گام برداری 13.....

۳.۲.۳ رتبه بندی 14.....

۲.۲.۳ تعریف 15.....

۱.۲.۳ قضیه 15.....

۲.۲.۳ قضیه 15.....

۲.۲.۳ مثال 15.....

۳.۳ سیستم های توزیع شده : مفاهیم و نتایج مقدماتی

۱.۳.۳ تعریف 17.....

۱.۳.۳ مثال 17.....

۱.۳.۳ کاربرد رتبه بندی گره های یک سیستم توزیع شده.....19

۲.۳.۳ مثال.....19

۳.۳.۳ مثال.....19

فصل ۱

مقدمه

شاید در اولین مواجهه با نظریه گراف سردرگم شوید و فکر کنید موضوعی ترسناک و نامفهوم باشد. اما در اصل نظریه گراف برخلاف تصور ما، چندان کارآمد نیست امت کاربردهای مفید مهمی در علوم مختلف ویا حتی زندگی روزمره دارد.

هدف از این پروژه یک مقدمه جامع در رابطه با نظریه گراف و درجه بندی راس های گراف است. در عصر مدرن امروزی، نظریه گراف به عنوان یک ابزار مفیدی برای تعیین کمیت و ساده سازی بسیاری از بخش های متحرک سیستم های پویا در نظر گرفته می شود و با توجه به مجموعه ای از گره ها و اتصالات، می توان هر چیزی را از طرح بندی شهر گرفته تا داده های کامپیوتری، به صورت یه زبان ریاضی معنادار تبدیل کرد. امروزه برنامه های کاربردی زیادی برپایه نظریه گراف طراحی و توسعه یافته است.

فصل ۲

در دنیای ما ، وضعیت های فراوانی وجود دارند که می توان توسط نموداری متشکل از یک مجموعه نقاط ، به علاوه خطوطی که برخی از این نقاط را به یکدیگر متصل میکنند ، به توصیف آن پرداخت . به عنوان مثال برای نشان دادن رابطه ی دوستی بین یک دسته از انسان ها می توانیم هر شخص را با یک نقطه مشخص کنیم و نقاط متناظر با هر دو دوست را با یک خط به یکدیگر وصل کنیم. یا در جای دیگر ممکن است برای نشان دادن یک شبکه ارتباطی ، از نموداری استفاده کنیم که در آن، نقاط نمایانگر مراکز ارتباطی و خطوط، نشان دهنده پیوند های ارتباطی بین مراکز باشند. توجه داشته باشید که در این گونه نمودار ها آنچه بیشتر مورد توجه است این است که آیا دو نقطه داده شده ، به وسیله یک خط به یکدیگر مستقل هستند یا نه و طریقه اتصال آنها اهمیتی ندارد.

تجرید ریاضی این وضعیت ها به مفهوم گراف منتهی می شود.

تعریف ۱.۱.۲: گراف : یک گراف از یک مجموعه نقاط نامتناهی راس ها $V(G)$ و یک مجموعه یال ها $E(G)$ که از راس ها مجزا هستند، تشکیل شده است.

یک سه تایی مرتب $(V(G), E(G), \psi(G))$ است که تشکیل شده از یک مجموعه ناتهی $V(G)$ از راس ها، یک مجموعه $E(G)$ مجزا از $V(G)$ از یال ها و یک تابع وقوع $\psi(G)$ که به هر یال G یک زوج نامرتب از راس های G را که الزاما متمایز نیستند نسبت میدهد. اگر e یک یال u و v دو راس باشند به طوری که $\psi_e = uv$ در این صورت گفته میشود که e ، راس های u و v را به یکدیگر وصل کرده است و راس های u, v دوسر یال e نامیده می شوند.

مثال ۱.۱.۲:

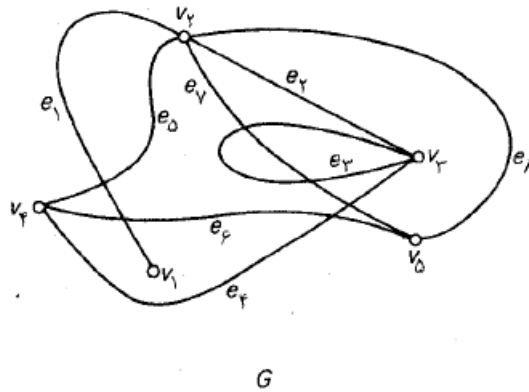
$$G = (V(G), E(G), \psi(G))$$

$$V(G) = \{V_1, V_2, V_3, V_4, V_5\}$$

$$E(G) = \{e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8\}$$

حال $\psi(G)$ به این صورت تعریف می شوند:

$$\begin{aligned} \psi_G(e_1) &= v_1v_2, & \psi_G(e_2) &= v_2v_3, & \psi_G(e_3) &= v_3v_3 \\ \psi_G(e_4) &= v_4v_3, & \psi_G(e_5) &= v_2v_4, & \psi_G(e_6) &= v_4v_5 \\ \psi_G(e_7) &= v_2v_5, & \psi_G(e_8) &= v_2v_5 \end{aligned}$$



مثال ۲.۱.۲:

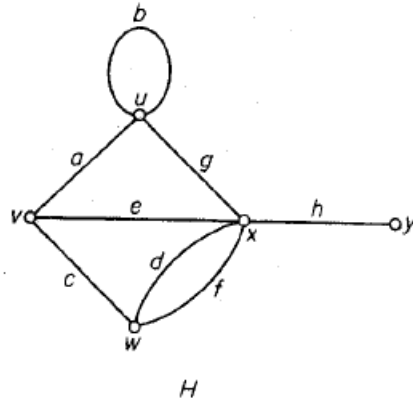
$$V(H) = \{u, v, w, x, y\}$$

$$E(G) = \{a, b, c, d, e, f, g, h\}$$

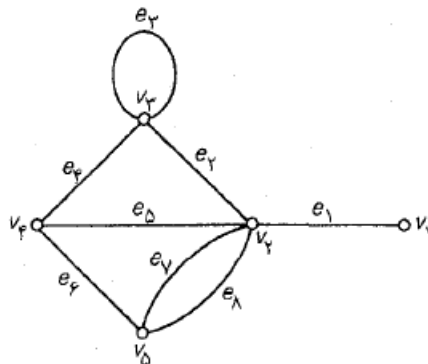
حال $\psi(G)$ به این صورت تعریف می شوند :

$$\psi_H(a)=uv , \psi_H(b)=uu , \psi_H(c)=uw , \psi_H(d)=wx$$

$$\psi_H(e)=vx , \psi_H(f)=wx , \psi_H(g)=ux , \psi_H(h)=xy$$



نکته : برای رسم یک گراف ، روش یکتایی وجود ندارد. بدین دلیل که موقعیت نسبی نقاط و خطوط که به ترتیب نمایانگر راس ها و یال های گراف هستند ، برای ما اهمیتی ندارد . به عنوان مثال نمودار دیگری از گراف در شکل آمده است.



تعریف ۲.۱.۲: طوقه : گرافی که به خودش وصل شود طوقه نام دارد. (یک یال با دوسر یکسان).

یال چندگانه: گرافی که از یک راس دوبار به راس دیگر رفته باشد یال چندگانه نامیده میشود.

تعریف ۳.۱.۲: گراف متناهی : اگر مجموعه راس ها و یال های یک گراف متناهی باشد ، گراف متناهی نامیده می شود.

تعریف ۴.۱.۲: گراف ساده : یک گراف ساده است ، اگر هیچ طوقه ای نداشته باشد و بین هر دو راس ان ، بیش از یک یال نباشد.

تعریف ۵.۱.۲: گراف مرکب : گرافی که ساده نباشد مرکب نامیده می شود. به عبارتی گرافی که دارای طوقه و یال چند گانه باشد گراف مرکب است.

تعریف ۶.۱.۲: درجه راس ها : درجه راس V گراف در G ، $d(V)$ برابر تعداد یال های واقع بر V می باشد. در این تعریف هر طوقه به عنوان دو یال شمرده میشود.

کمترین و بیشترین درجه راس های G را به ترتیب با $\delta(G)$ و $\Delta(G)$ نمایش می دهیم.

قضیه ۱.۲.۳: $\sum d(v) = 2E$

نتیجه: در هر گراف تعداد راس های فرد ، زوج است.

فصل ۳

1.3: سرآغاز و انگیزه:

در شبکه‌ها (گراف) به ویژه در شبکه‌های جهانی وب، شبکه‌های اجتماعی و شبکه‌های زیستی و همچنین در سیستم‌های توزیع شده، به منظور تعیین اهمیت گره‌ها، از معیارهای مرکزیت استفاده می‌شود. یک معیار مرکزیت، تابعی است که به هر کدام از گره‌های شبکه متناسب با اهمیت آن گره، عددی حقیقی نسبت می‌دهد.

معیارهای مرکزیت را بر اساس روش‌های گوناگون از قبل مرکزیت مبتنی بر درجه، بردار ویژه، رتبه بندی، و نزدیکی تعریف می‌کنند. در ایم میان روش رتبه بندی که برای اولین بار توسط برین و پیچ، به منظور تعیین اهمیت صفحه‌های وب و در حوزه الگوریتم‌های وب مطرح شد و موجب شکل‌گیری شرکت‌های گوگل گردید. به دلیل کاربردهای گسترده‌ای که بعداً در سایر حوزه‌های علوم کامپیوتر پیدا کرد از اهمیت ویژه‌ای برخوردار است.

اکنون چند مثال را به منظور بررسی اهمیت رتبه بندی گره‌های یک شبکه در مسایل کاربردی دنبال می‌کنیم.

مثال ۱.۱.۳: (کاربرد در شبکه وب)

اگر صفحه‌های وب را معادل با گره‌های شبکه و آبرلینک‌های را که بین صفحه‌های وب پیوند ایجاد می‌کنند، لینک‌های شبکه در نظر بگیریم، شبکه ایجاد شده را شبکه وب می‌نامیم.

پیمایش کنندگان صفحه های وب با جستجوی یک عبارت در موتورهای جستجو مانند گوگل یا بینگ ، غالباً با صدها یا هزاران صفحه مواجه میشوند که بررسی همه این صفحه ها زمان بر و در بسیاری اوقات ناممکن است لذا باید صفحه های مذکور به شیوه ای مناسب مرتب شوند و صفحه هایی که اهمیت بیشتر دارند در ابتدای نتیجه جستجو قرار بگیرند.

مثال ۲.۱.۳: (کاربرد در امور نظامی)

اگر سربازان و فرماندهان یک لشکر نظامی را گره های شبکه ارتباطی بین آنها را لینکهای شبکه در نظر بگیریم ، شبکه ایجاد شده را یک شبکه نظامی می نامیم. به منظور انجام حمله نظامی و رساندن بیشترین آسیب ممکن به دشمن میتوانیم همه نیروی های دشمن را مورد حمله قرار دهیم . واضح است که این عملیات پرهزینه و زمان بر خواهد بود. از آنجا که نابودی گره های با اهمیت فرماندهان باعث از هم گسیختگی شبکه لشکر خواهد شد، لذا به جای حمله به همه گره های شبکه میتوانیم تنها به گره های مهم حمله کنیم برای این کار نیاز است که نخست گره های مهم را شناسایی کنیم.

تعریف ۱.۱.۳: رتبه بندی : رتبه بندی مبتنی بر یک زنجیره مارکوف است که در حالت متمرکز، با به توان رساندن ماتریس انتقال زنجیر مارکوف محاسبه می شود.

2.3 مفاهیم و نتایج مقدماتی

در این بخش معیار مرکزیت ، معیار مرکزیت مبتنی بر درجه ، زنجیر مارکوف ، گام برداری تصادفی، رتبه بندی و الگوریتم متمرکز محاسبه رتبه بندی را توضیح می دهیم.

مرکزیت ۱.۲.۳: به گره هایی از شبکه که دارای اهمیت بیشتری هستند،

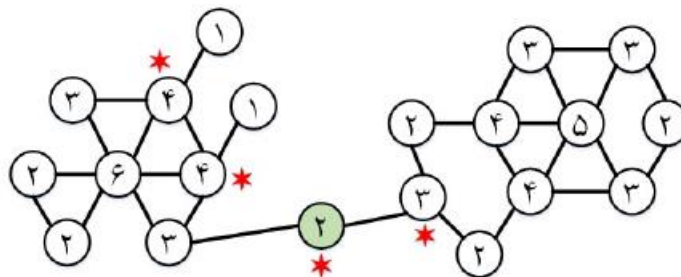
گره ها مرکزی میگوئیم. در واقع مفهوم مرکزیت معادل با این پرسش است : « کدام گره ها ، شبکه مرکزی هستند؟ »

تعریف ۱.۲.۳: فرض کنید $G=(V,E)$ یک گراف باشد. هر تابع $C:V \rightarrow R$ یک معیار مرکزی C نامیده می شود که در آن، $C(V)$ اهمیت گره $V \in V$ را نشان می دهد.

برای تعیین مرکزیت، معیارها متفاوتی ارائه شده است و یکی از رایج ترین معیارهای مرکزی معیار مبتنی بر درجه است که در آن، اهمیت هر گره متناسب با درجه آن گره است.

در شبکه ها گره هایی که درجه بیشتری دارند، اهمیت بیشتری برخوردارند. اما عکس این موضوع صادق نیست. به بیان دیگر گره هایی که با درجه کوچک هستند الزاماً کم ارزش نیستند.

مثال ۱.۲.۳: در شکل ۱ با اینکه درجه گره خاکستری کوچک است ولی چون حذف آن موجب ناهبندی شبکه می شود، از اهمیت زیادی برخوردار است، رأس هایی مانند راس های ستاره دار در (شکل ۱) که حذف آنها موجب ناهبندی شدن گراف می شود، نقاط مفصلی یا راس های برشی می گوئیم.



زنجیر مارکوف و گام برداری تصادفی ۲.۲.۳:

یک فرآیند تصادفی خانواده ای از متغیرهای تصادفی مانند $X=\{X_t:t \in T\}$ با مجموعه اندیس گذار T فضای حالت S این مورد ها در نظر گرفته می شود:

$$T=[0, \infty) \text{ یا } T=\{0,1,2,\dots\} \text{ (الف)}$$

$$S=N \text{ یا } S=R \text{ (ب)}$$

به این ترتیب، در یک تقسیم بندی کلی، فرآیندهای تصادفی به چهار تقسیم می شوند: زمان گسسته - حالت گسسته، زمان گسسته- حالت پیوسته، زمان پیوسته- حالت گسسته و زمان پیوسته - حالت پیوسته. یک فرآیند تصادفی زمان گسسته - حالت گسسته را زنجیر مارکوف می نامیم هرگاه به ازای $n \in \mathbb{N}$ و هر $i_0, i_1, \dots, i_{n-1}, j = s$ داشته باشیم.

$$P(X_{n+1}=j | X_0=i_0, \dots, X_{n-1}=i_{n-1}, X_n=i) = P(X_{n+1}=j | X_n=i)$$

رتبه بندی ۳.۲.۳: روش رتبه بندی در ابتدا برای تعیین اهمیت به صفحه های وب مطرح شد. یک پیمایش کننده صفحه های وب را در حال مشاهده صفحه S در وب A در نظر بگیرید.

پس از مشاهده A یکی از چهار پیشامد زیر ممکن است رخ بدهد:

- 1) پیمایش کننده به پیمایش صفحه های وب پایان می دهد.
- 2) پیمایش کننده از طریق یکی از ابر لینک ها خروجی A وارد قسمت دیگر A میشود.
- 3) پیمایش کننده از طریق یکی از ابر لینک ها خروجی A وارد قسمت B می شود.
- 4) پیمایش کننده وارد صفحه C میشود که جز لینک های خروجی A نیست.

برای محاسبه رتبه بندی شبکه وب را توسط گراف جهت دار $D=(V,A)$ که در آن A, V به ترتیب مجموعه راس ها و یال ها هستند مدل سازی میکنیم:

در گراف D یک تناظر یک به یک بین راس ها و صفحه های وب، برقرار است و از راس U به راس V یک یال وجود دارد اگر و تنها اگر صفحه متناظر با راس های U دارای یک ابر لینک خروجی به صفحه متناظر با راس V باشد. فرض کنیم $n = |V|$ برای هر گراف D ماتریس مربعی $P_{n \times n}$ به صورت زیر می سازیم:

$$\tilde{p}_{v,u} = \begin{cases} \begin{cases} \frac{1}{|V_O(v)|} & u \in V_O(v) \\ 0 & \text{اگر نه} \end{cases} & |V_O(v)| > 0 \\ \frac{1}{n} & |V_O(v)| = 0 \end{cases}$$

از ماتریس P برای ساخت ماتریس انتقال یک زنجیر مارکوف استفاده می کنیم.

تعریف ۲.۲.۳: فرض کنید $D=(V,A)$ یک گراف جهتدار باشد. توزیع مانای زنجیر مارکوفی را که فضای حالت آن مجموعه V و ماتریس انتقال آن

$$P^* = \varepsilon P + (1 + \varepsilon) \left(\frac{1}{n}\right) J$$

باشد ، رتبه بندی گویند.

در اینجا J یک ماتریس مربعی $n \times n$ با درایه های 1 و ε عددی حقیقی در بازه $(0,1)$ است که احتمال باقی ماندن گره را نشان می دهد.

قضیه ۱.۲.۳: فرض کنید P ماتریس انتقال یک زنجیر مارکوف باشد ، اگر P نامتناوب و تتحویل پذیر باشد ، انگاه یک بردار سطری π با شرایط $\pi P = \pi$ و $\pi I = 1$ که در آن I یک بردار ستونی است که همه درایه های آن ، I هستند .

به عبارت دیگر $\lim_{t \rightarrow \infty} P^t \geq 0$ که از J مستقل هستند .

قضیه ۲.۲.۳: ماتریس انتقال P در تعریف ۲.۲.۳ و در شرایط قضیه ۱.۲.۳ صادق است.

مثال ۲.۲.۳:

قرار می دهیم $\varepsilon = .85$.

ماتریس P را به این صورت می نویسیم:

$$\tilde{P} = \begin{bmatrix} 0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{bmatrix} .$$

برای ساختن P براساس تعریف به صورت زیر عمل می کنیم:

$$P = \epsilon \tilde{P} + (1 - \epsilon)(1/n)J = \begin{bmatrix} 0,030 & 0,455 & 0,030 & 0,455 & 0,030 \\ 0,030 & 0,030 & 0,880 & 0,030 & 0,030 \\ 0,880 & 0,030 & 0,030 & 0,030 & 0,030 \\ 0,030 & 0,030 & 0,030 & 0,030 & 0,880 \\ 0,200 & 0,200 & 0,200 & 0,200 & 0,200 \end{bmatrix}$$

برای محاسبه بردار مانا (بردار رتبه بندی) ، معادله $\pi P = \pi$ را تشکیل می دهیم که از آن دستگاه معادلات خطی به دست می آید:

$$\begin{cases} 0,030x_1 + 0,030x_2 + 0,880x_3 + 0,030x_4 + 0,200x_5 = x_1 \\ 0,455x_1 + 0,030x_2 + 0,030x_3 + 0,030x_4 + 0,200x_5 = x_2 \\ 0,030x_1 + 0,880x_2 + 0,030x_3 + 0,030x_4 + 0,200x_5 = x_3 \\ 0,455x_1 + 0,030x_2 + 0,030x_3 + 0,030x_4 + 0,200x_5 = x_4 \\ 0,030x_1 + 0,030x_2 + 0,030x_3 + 0,880x_4 + 0,200x_5 = x_5 \end{cases}$$

با حل معادلات دستگاه بالا به شرط $\pi I = 1$ بردار مانا به این صورت می باشد :

$$\pi \approx [0,243 \quad 0,169 \quad 0,209 \quad 0,169 \quad 0,209]$$

3.3 سیستم های توزیع شده : مفاهیم و نتایج مقدماتی

در بخش قبل ، روش محاسبه متمرکز رتبه بندی را بررسی کردیم . چون الگوریتم روی یک سیستم متمرکز اجرا می شود ، آن را متمرکز می نامیم. در مقابل سیستم های متمرکز سیستم های توزیع شده قرار دارند. الگوریتم رتبه بندی بر روی سیستم های توزیع شده نیز قابل محاسبه است.

تعریف ۱.۳.۳: یک سیستم توزیع شده مجموعه ای از گره های محاسباتی است ، که توسط یک شبکه ارتباطی به هم وصل شده اند و در صدد انجام کاری مشترک هستند.

منظور از گره محاسباتی هر وسیله ای است که قابلیت پردازش و نگهداری داده ها را داشته باشد. لذا رایانه های شخصی ، حسگرهای یک شبکه حسگر بیسیم ، پردازنده های یک سیستم چند پردازنده ای و یا تلفن های همراه ، گره های محاسباتی یک سیستم توزیع شده هستند و سیستم های توزیع شده واحد های محاسباتی در صدد انجام کاری مشترک هستند.

مثال ۱.۳.۳: گراف G را به صورت شکل ۳ در نظر می گیریم. که در آن هر گره معادل با یک واحد محاسباتی در یک سیستم توزیع شده است. هدف ما مجموع اعداد فهرست

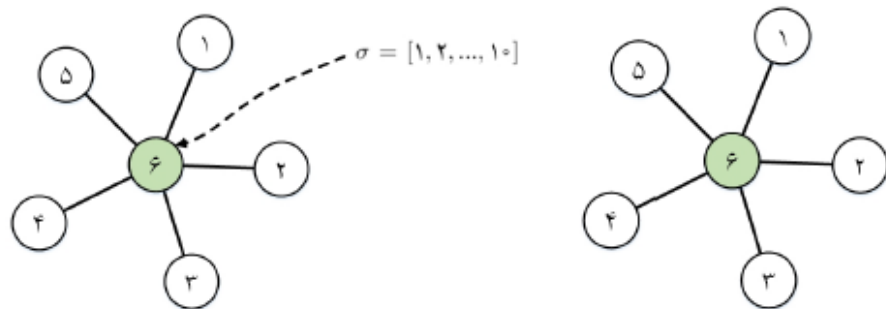
$\sigma = [1,2,\dots,10]$ است که می خواهیم محاسبه کنیم. پس σ را در سیستم توزیع شده ، توزیع و سپس مجموع آن را حساب می کنیم.

فرض کنید σ بر روی رایانه شماره 6 قرار داشته باشد به جای اینکه کل محاسبات را توسط رایانه 6 انجام دهیم ، رایانه 6 فهرست σ را به 5 زیر فهرست می شکند. (چون 5 واحد محاسبات داریم).

و هر زیرفهرست را به یکی از واحد های محاسباتی دیگر ارسال می کند.

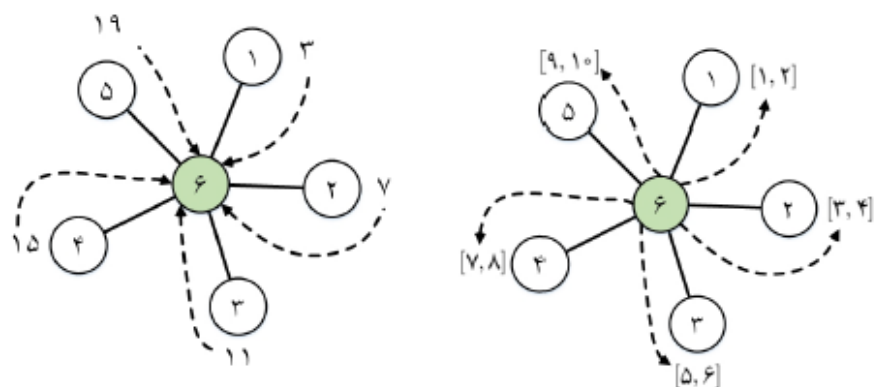
در ادامه گره های 1 و 2 و 3 و 4 و 5 مجموع اعداد فهرست های خودشان را حساب و نتیجه را به گره 6 ارسال می کنند.

در پایان گره 6 مجموع عددهای رسیده از گره های دیگر را به عنوان خروجی اعلام می کند و کار پایان می یابد.



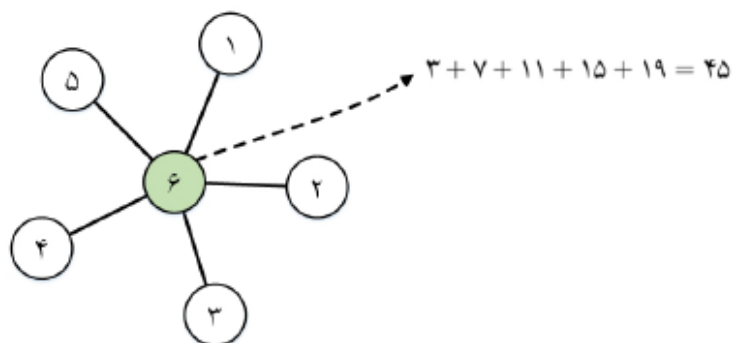
(ب) فهرست σ برای انجام محاسبات به گره ۶ ارسال می‌شود.

(آ) گراف G



(ج) گره ۶، فهرست σ را به چند زیرفهرست می‌شکند و هر زیرفهرست را به یک واحد محاسباتی دیگر ارسال می‌کند.

(د) هر گره مجموع اعداد فهرست خود را محاسبه می‌کند و نتیجه را برای گره ۶ ارسال می‌کند.



کاربرد رتبه بندی گره های یک سیستم توزیع شده ۱.۳.۳ : یکی از مسایل مطرح در حوزه سیستم های توزیع شده ، توازن بار است.

اگر در انجام محاسبات توزیع شده ، بار محاسباتی به گونه ای متوازن بین واحد های محاسباتی توزیع شود ، گوئیم توازن بار روی داده است .

در مثال قبل بار محاسباتی، زیرفهرست هایی است که توسط گره 6 به گره های 1 تا 5 تخصیص داده شده است و اگر طول فهرست ها برابر یا تقریبا برابر باشد ، توازن بار روی داده است.

هدف از توازن بار ، استفاده بهینه از منابع ، کاهش زمان پاسخ و از بین بردن سر بار محاسباتی واحدهای محاسباتی است.

مثال ۲.۳.۳: (کاربرد مسیریابی)

ساختار و توپولوژی بسیاری از شبکه ها دایما در حال تغیر است. لذا برای ارسال پیام از یک گره به گره دیگر ، لازم است که مسیریابی به صورت پویا انجام شود.

از انجا که گره های با رتبه ای بالا، با تعداد زیادی از گره های دیگر در ارتباط هستند ، یک گره برای اینکه پیامی را انتقال دهد، از بین همسایه های خود به همسایه های با رتبه بالاتر اولویت بیشتر می دهد و از این طریق موجب افزایش سرعت رسیدن پیام به مقصد می شود.

مثال ۳.۳.۳: (کاربرد در زمان بندی کارها)

چون در سیستم رایانه ای در یک بازه زمانی کوتاه نیاز به پردازش های بسیار است ، اولویت بندی کارها در این سیستم ها امری ضروری است. چارچوب هدوپ، به منظور ذخیره و پردازش توزیع شده به کار می رود. یکی از روش های نوین برای اولویت بندی کارها در چارچوب هدوپ، استفاده از رتبه کارها است.