دسترسی در سایتِ http://jnrm.srbiau.ac.ir
سال پنجم، شماره بیستم، مهر و آبان ۱۳۹۸
شماره شاپا: ۵۸۸X–۲۵۸۸

دانشگاه آزاد اسلامی، واحد علوم و تحقیقات
پژوهش‌های نوین در ریاضی

مدل‌بندی مسأله هضم جزئی بصورت مساله شبکه جریان

رضا ندیمی[۱*]، امید رنجبر[۲]

(۱) استادیار، گروه علوم کامپیوتر، دانشکده علوم ریاضی، دانشگاه مازندران، بابلسر، ایران
(۲) گروه علوم کامپیوتر، دانشکده علوم ریاضی، دانشگاه مازندران، بابلسر، ایران

تاریخ ارسال مقاله: ۹۶/۰۷/۰۶ تاریخ پذیرش مقاله: ۹۸/۰۴/۲۲

چکیده

نگاشت نقاط مرزی[۱] یکی از مسائل جالب توجه در زیست‌شناسی محاسباتی به شمار می‌رود. یک رشته DNA به صورت دنباله‌ای از حروف A, T, C, G می‌باشد. هنگامی که یک آنزیم محدود کننده به یک محلول DNA اضافه می‌شود، مولکول DNA از مکان‌های خاصی بریده می‌شود. هدف از نگاشت نقاط مرزی پیدا کردن نقاط برش برای یک آنزیم معین است. در روش هضم جزئی، برش‌ها طوری انجام می‌شود که فاصله دو به دوی همه نقاط برش حاصل شود. در بیان ریاضی مساله، فاصله دو به دوی n نقطه واقع بر یک پاره خط داده شده است و هدف بدست آوردن خود این نقاط است. در بیوانفورماتیک این مساله به مساله هضم جزئی[۲] معروف شده است. در این مقاله یک مدل شبکه جریان تعمیم یافته برای مساله ارایه می‌دهیم. با توجه به اینکه کلاس پیچیدگی این مساله یکی از قدیمی‌ترین و مهمترین مسایل باز در بیوانفورماتیک نظری است (تاکنون نه الگوریتمی با زمان چند جمله‌ای و نه اثباتی بر Np-complete بودن آن ارایه شده است)، کاهش مساله PDP به مساله شبکه جریان دریچه جدیدی را برای چالش با این مساله می‌گشاید.

واژه‌های کلیدی: نگاشت نقاط مرزی، شبکه جریان، مسأله هضم جزئی، توالی DNA.

---

*. عهده‌دار مکاتبات:
Email: nadimi@umz.ac.ir

1. Restriction site mapping
2. Partial Digest Problem

۱- مقدمه

یکی از مسائل جالب توجه در زیست‌شناسی محاسباتی، نگاشت نقاط مرزی است. هنگامی که یک آنزیم محدود کننده خاص به یک محلول DNA اضافه شود، DNA از مکان‌های با توالی مشخصی بریده می‌شود. هدف از نگاشت نقاط مرزی، محاسبه‌ی موقعیت محل‌های برش برای یک آنزیم مشخص است. با استفاده از برخی از آزمایشات بیوشیمیایی می‌توان فاصله‌ی بین هر جفت از مکان‌های مرزی را پیدا کرد. در مسأله‌ی PDP فاصله‌های بدست آمده از عمل هضم توسط یک آنزیم به ما داده شده است (شکل ۱) و هدف محاسبه‌ی همه‌ی محل‌های برش است.

فرض کنید $X = \{x_0. x_1. \dots. x_n\}$ مجموعه‌ی محل‌های برش توسط یک آنزیم خاص روی یک رشته DNA باشد. تعداد $N = \binom{n+1}{2}$ فاصله‌ی دو به دو بین این مکان‌ها را با چند مجموعه‌ای ΔX نشان می‌دهیم:

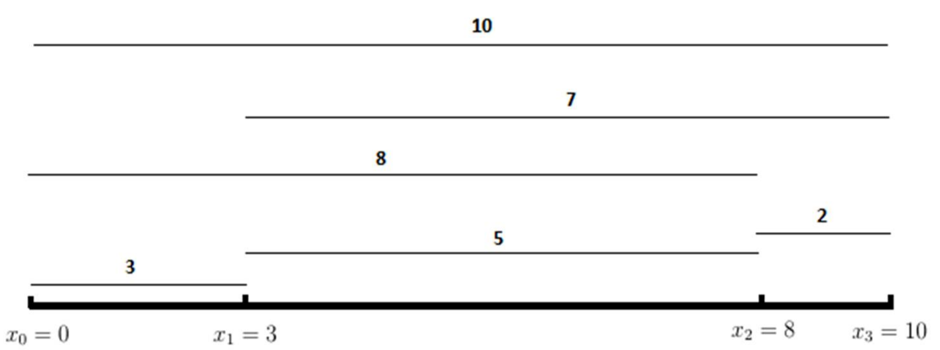
$$\Delta X = \{x_j - x_i | x_j > x_i;\ i.j = 0.1. \dots. n\} \quad (۱)$$

در مسأله‌ی PDP چندمجموعه‌ای $B = \{b_1. b_2. \dots. b_N\}$ از فاصله‌ها داده شده است، هدف یافتن مجموعه‌ی $Y = \{y_0. y_1. y_2. \dots. y_n\}$ از نقاط واقع بر روی یک خط است به‌طوری که فاصله‌های دو به دوی این نقاط از هم برابر چندمجموعه‌ای B باشد.

در این مقاله یک مدل شبکه‌ی جریان برای حل مسأله‌ی PDP ارائه می‌کنیم. در ادامه، در بخش دوم پیشینه تحقیق را مرور خواهیم کرد. در بخش سوم مدل جدید را به فرم یک مسأله‌ی شبکه جریان معرفی می‌کنیم. در بخش چهارم با تبدیل مساله شدنی بودن به مساله حداکثر جریان آن را حل می‌کنیم و در بخش پنجم به اثبات صحت مدل خواهیم پرداخت.

۲- پیشینه تحقیق

این مسأله در دهه‌ی ۱۹۳۰ در حوزه‌ی کریستالوگرافی اشعه X مطرح شد[۶]. در سال ۱۹۸۸ P.Lemke و M.Werman این مسأله را در زمان شبه چندجمله‌ای حل کردند (زمان اجرای الگوریتم ارائه شده‌ی آنها به بزرگترین مقدار عضو B وابسته بود)[۵]. Skiena و همکارانش یک الگوریتم عقبگرد برای حل این مسأله ابداع کردند که زمان اجرای آن تنها به اندازه ورودی وابسته بود [۷]. در سال ۱۹۹۴ Z.Zhang با استفاده از یک مثال، نشان داد که زمان اجرای الگوریتم عقب‌گرد در بدترین حالت نمایی است[۸]. T. Dakice در رساله‌ی دکترای خود یک مدل برنامه‌نویسی صفر و یک درجه دوم برای مسأله‌ی PDP ارائه کرد و آن را با استفاده از یک الگوریتم مکاشفه‌ای برنامه ریزی نیمه معین متوالی حل کرد[۴]. صالحی فتح‌آبادی و ندیمی یک مدل پیوسته برای این مسأله در سال ۲۰۱۰ ارائه کردند[۹]. ندیمی و همکارانش در سال ۲۰۱۱ الگوریتم سریعی برای حل این مساله ارائه کردند که عملکرد بهتری نسبت به الگوریتم‌های پیشین دارد[۱۰].



شکل ۱. قطعه‌های بریده شده روی مولکول DNA

۳- مدل شبکه جریان برای مسأله PDP

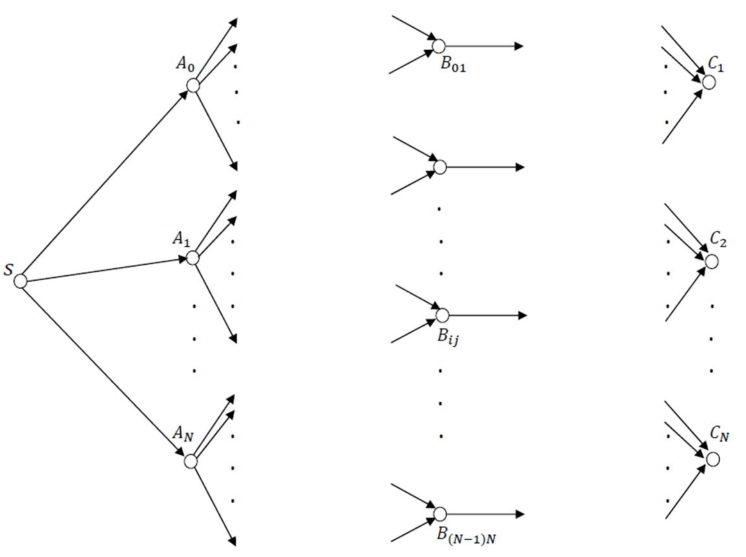
در بخش قبل دیدیم که در مسأله‌ی PDP، چندمجموعه‌ای $B = \{b_1 . b_2 . \dots . b_N\}$ از فاصله‌ها داده شده است و هدف پیدا کردن مجموعه‌ی $Y = \{y_. y_1 . y_2 . \dots . y_n\}$ از نقاط واقع بر یک خط است به طوری که فاصله‌های دوبه دوی این نقاط از هم برابر چند-مجموعه‌ای B باشد. از آنجایی که با اضافه یا کم کردن یک مقدار ثابت به همه اعضای مجموعه‌ی Y، چند-مجموعه‌ای حاصل از فاصله‌های دوبه دوی آنها تغییر نمی‌کند، بدون از دست دادن کلیت مساله می‌توان با انتقال همه اعضای Y، y_0 را برابر صفر در نظر گرفت. بدین ترتیب واضح است که $y_1 . y_2 . \dots . y_n$ در چندمجموعه‌ای B وجود خواهند داشت و y_n برابر بزرگترین عضو Y خواهد بود. بنابراین مسأله‌ی PDP معادل است با پیدا کردن n عضو از B به‌طوری که فاصله‌ی دوبه دوی مجموعه‌ی حاصل از این نقاط به همراه نقطه صفر، برابر B باشد. به منظور پیدا کردن این n عضو از B، یک مدل شبکه جریان طراحی می‌کنیم که هر جریان شدنی برای این شبکه مشخص کننده‌ی این نقاط باشد.

در (شکل ۲) یک مدل شبکه جریان "صفر و یک" با چهار لایه از گره‌ها و سه لایه از یال‌ها می‌سازیم.

لایه‌ی اول از گره‌ها تنها شامل یک گره‌ی شروع S است. در لایه دوم $N+1$ گره داریم که آنها را با $A_0 . A_1 . \dots . A_N$ نشان می‌دهیم. A_0 را متناظر با $b_0 = 0$ در نظر می‌گیریم. همچنین گره‌های $A_1 . A_2 . \dots . A_N$ را به $b_1 . b_2 . \dots . b_N$ اختصاص می‌دهیم. برای اعضای تکراری می‌توان در این لایه فقط یک گره در نظر گرفت که در اینصورت اندازه شبکه کوچکتر خواهد شد. ما در اینجا همه $N+1$ گره را در نظر گرفته‌ایم.

در لایه‌ی سوم گره‌ها، برای هر جفت $i < j$ یک گره به نام B_{ij} در نظر می‌گیریم اگر و فقط اگر داشته باشیم $|b_i - b_j| \in B$. در این لایه حداکثر $\frac{N(N+1)}{2}$ گره وجود دارد.

در لایه‌ی آخر از گره‌ها، $|B| = N = \binom{n+1}{2}$ گره وجود دارد که متناظر b_iها هستند. گره‌ها در این لایه را با $C_1 . C_2 . \dots . C_N$ نمایش می‌دهیم. در صورت وجود اعداد تکراری در B برای هریک گره‌ای مجزا در لایه آخر در نظر گرفته می‌شود. حال به یال‌های شبکه می‌پردازیم.



شکل۲. مدل شبکه جریان (G)

لایه‌ی اول یال‌ها حداکثر شامل $N+1$ یال از گره‌ی مبدا به A_iها است. در این لایه، هر یال یک ضریب با مقدار n دارد. به این معنی که اگر $f(S.\ A_i)$ مقدار جریان ورودی به یال $(S.\ A_i)$ باشد، $n \times f(S.\ A_i)$ مقدار جریانی است که گره‌ی A_i دریافت خواهد کرد.

در لایه‌ی دوم یال‌ها، دو یال ورودی به هر گره‌ی B_{ij} خواهیم داشت، یک یال از A_i و یک یال از A_j. همه‌ی یال‌های این لایه ضریب $\frac{1}{2}$ دارند.

در لایه آخر، یک یال از هر گره‌ی B_{ij} خارج می‌شود. از گره‌ی B_{ij} به گره‌ی C_k یک یال داریم اگر و فقط اگر $b_k = |b_i - b_j|$. در این لایه هیچ ضریبی روی یال‌ها وجود ندارد (برابر ۱ است).

در این شبکه متغیرهای $f(u.v)$ به ازای همه کمان‌های $(u.v)$ متغیرهای "صفر و یک" هستند. در گره‌ی S به مقدار $n+1$ واحد عرضه‌ی جریان و در گره‌های C_i به مقدار -1 واحد تقاضا وجود دارد. در گره‌های لایه‌ی دوم و سوم هیچ عرضه و تقاضایی وجود ندارد. لازم به ذکر است که با در نظر گرفتن ضرایب افزایش و کاهش لایه اول و دوم کمان‌ها، مجموع عرضه دقیقاً به مقدار تقاضا افزایش می‌یابد. زیرا $(n+1)$ واحد جریان ارسالی از مبدا بعد از عبور از لایه اول با ضریب n به مقدار $n(n+1)$ واحد جریان تبدیل می‌شود. بعد از عبور از لایه دوم که همه کمان‌های آن دارای ضریب $\frac{1}{2}$ هستند، مقدار جریان به $n(n+1)/2$ کاهش می‌یابد که دقیقاً برابر مجموع تقاضا در لایه آخر یعنی $N = n(n+1)/2$ است.

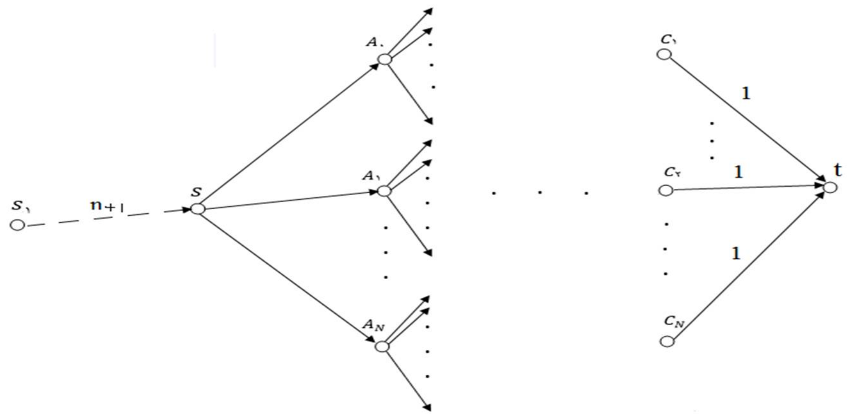
در ادامه، ابتدا با تبدیل مساله شدنی بودن به مساله حداکثر جریان، روشی برای حل آن ارایه می‌دهیم و سپس ثابت می‌کنیم که هر جواب شدنی برای این شبکه یک جواب برای مساله PDP را به دست می‌دهد.

۴- حل مسأله شدنی بودن

برای پیدا کردن یک جواب شدنی، به مدل شبکه جریان ساخته شده یک گره‌ی انتهایی t یا چاه اضافه می‌کنیم، گره‌ها در لایه‌ی چهارم به گره‌ی چاه متصل می‌شوند با توجه به مقدار تقاضا در گره‌های سطح آخر، از همه‌ی گره‌ها کمانی با ظرفیت ۱ به t وارد می‌شود. همچنین گره s_1 را مطابق شکل ۳ به شبکه اضافه کرده و یالی با ظرفیت $n+1$ از آن به s رسم می‌کنیم. ظرفیت این یال را برابر با میزان عرضه در s یعنی $n+1$ قرار می‌دهیم. به اینصورت شبکه جریان جدید G_1 مطابق شکل ۳ حاصل می‌شود.

به منظور یافتن جریان شدنی در شبکه G، کافی است مسأله‌ی حداکثر جریان[1] را در شبکه G_1 از گره‌ی منبع s_1 به چاه t حل کنیم. نشان می‌دهیم که اگر حداکثر جریان در شبکه G_1 تمامی یال‌های متصل به منبع و چاه را اشباع[2] کند، در این صورت می‌توان گفت که مساله اصلی یک جواب شدنی دارد، در غیر این صورت شدنی نخواهد بود. منظور از اشباع شدن یک یال عبور حداکثر جریان ممکن (برابر ظرفیت کمان) ازآن یال است. زمانی که در شبکه‌ی G_1 یال خروجی از منبع اشباع می‌شود، مقدار $n+1$ واحد جریان به گره S می‌رسدکه دقیقاً برابر مقدار عرضه این گره در شبکه G است. همچنین اشباع شدن یال‌های ورودی به چاه به این معنی است که روی هر یال ورودی به چاه، ۱ واحد جریان داشته باشیم. در این صورت در واقع از هر گره لایه آخر، یک واحد جریان به چاه وارد شده است که دقیقا برابر مقدار تقاضا در گره‌های لایه آخر G است. نتیجه اینکه شبکه G شدنی است اگر و فقط اگر در G_1 به ازای حداکثر جریان، یال‌های منبع و چاه اشباع شوند. با توجه به توضیح فوق، در صورت شدنی بودن شبکه G، مقدار حداکثر جریان در G_1 برابر $n+1$ خواهد بود. به همین صورت نتیجه می‌شود که اگر مقدار حداکثر جریان در G_1 کمتر از $n+1$ باشد، به مفهوم این است که G دارای جواب شدنی نمی‌باشد. پس با بدست آوردن حداکثر جریان در شبکه G_1، در صورتیکه مقدار جریان بیشینه برابر $n+1$ باشد، با حذف گره‌های s_1 و t، به یک جریان شدنی در G می‌رسیم.

1. Maximum Flow
2. Saturate



شکل۳. مساله ماکزیم جریان (شبکه G_1)

در اینجا مدل ریاضی مساله حداکثر جریان شبکه فوق را ارایه می‌دهیم.

الف – تابع هدف:

$$Maximize \quad z$$

ب – محدودیت ظرفیت یال‌ها:

$$0 \le f(u,v) \le c(u,v)$$

در اینجا $c(u.v)$ ظرفیت جریان ورودی به کمان $(u.v)$ است. در کمان خروجی از مبدا، این ظرفیت برابر $n+1$ و در بقیه کمان‌ها برابر ۱ است.

ج – شرط بقای جریان:

برای تمام $u \in V-\{s,t\}$ شرط بقای جریان برای تابع جریان f به صورت زیر است:

$$\sum_{v\in V} \alpha(v.u) f(v.u) = \sum_{v\in V} f(u.v)$$

که در آن $\alpha(u.v)$ ضریب افزایش (کاهش) روی کمان‌هاست. همچنین برای گره مبدا داریم:

$$\sum_{v\in V} f(s_1.v) = z$$

با توجه به شرط بقای جریان در گره‌های میانی و تعمیم یافته بودن جریان، قیدی برای گره مقصد لازم نیست زیرا هر جریان خروجی از مبدا بعد از افزایش و کاهش توسط ضرایب کمان‌ها در نهایت به مقصد خواهد رسید.

شرط صفر و یک را برای جریان همه کمان‌ها بجز کمان $(s_1.s)$ به صورت زیر در نظر می‌گیریم:

$$\forall (u.v) \neq (s_1.s): \quad f(u.v) = 0 \ or\ 1$$

۵- تبدیل جریان شدنی به جواب مسأله‌ی PDP

نشان می‌دهیم هر جریان شدنی در شبکه جریان G، یک راه حل برای مسأله‌ی PDP است. جواب شدنی در G جریانی است که $n+1$ واحد عرضه را با اعمال ضرایب موجود به t انتقال دهد به‌طوری که شرط بقای جریان برای گره‌های لایه‌های میانی رعایت شود.

با توجه به صفر و یک بودن شبکه G، باید $n+1$ گره از گره‌های A_i برای دریافت مقدار جریان موجود در S انتخاب شوند. گره‌های انتخاب شده بی شک شامل A_0 و A_N خواهند بود زیرا تنها راه فرستادن جریان به C_N استفاده همزمان از این دو گره است. نشان می‌دهیم اعضایی از چندمجموعه‌ای B که متناظر گره‌های انتخاب شده از A_iها هستند تشکیل یک جواب برای مساله PDP می‌دهند. با توجه به ضریب روی یال‌های خروجی گره S، به هر یک از گره‌های انتخاب شده از لایه دوم مقدار $n \times 1 = n$ واحد جریان وارد می‌شود. به دلیل صفر و یک بودن متغیرها، هر یک از گره‌های انتخاب شده از لایه دوم، جریان دریافتی را توسط n یال خروجی به n گره از لایه سوم می‌فرستند. گره های لایه سوم دارای یک یال خروجی با ضریب ۱ و دو یال ورودی با ضریب $\frac{1}{2}$ هستند، بنابراین یا هیچ جریانی از این گره‌ها

عبور نمی‌کند ویا هر سه یال اشباع می‌شوند. به همین دلیل زمانی که یکی از گره‌های B_{ij} واسط جریان است، اولاً تقاضای دقیقاً یک گره از لایه آخر متناظر با یک b_k را برآورده می‌کند و ثانیاً با توجه به اینکه در ساخت G کمان‌های ورودی B_{ij}ها از گره‌های A_i و A_jای انتخاب شدند که متناظر اعداد b_i و b_j بودند و در شرط $|b_i - b_j| = b_k$ صدق می‌کردند، تفاضل b_iهای دو گره انتخاب شده از لایه دوم دقیقاً برابر b_k است. بنابراین وقتی یک جواب شدنی داریم و همه تقاضاهای گره‌های لایه آخر برآورده شده‌اند، می‌توان گفت تفاضل دو بدوی اعضایی از B که متناظر گره‌های انتخاب شده لایه دوم هستند، دقیقاً برابر خود چندمجموعه‌ای B است. یعنی گره‌های انتخاب شده از لایه دوم توسط یک جریان شدنی، اعضایی از B را مشخص می‌کنند که جواب مساله PDP هستند.

مثال: در اینجا برای درک بهتر مدل، یک نمونه از مساله‌ی PDP را به همراه دو جواب ممکن از آن ارایه می‌دهیم. برای هر دو جواب PDP، جریان‌های شدنی معادل را در مدل شبکه جریان ارایه می‌دهیم.

ورودی مساله (در اینجا: n=3 و N=6):

$B = \{2.3.5.7.8.10\}$

یکی از جواب‌ها:

$X = \{x_0 = 0. x_1 = 3. x_2 = 8. x_3 = 10\}$

مدل شبکه جریان و جریان شدنی متناظر جواب در شکل ۴ آورده شده‌اند.

کمان‌هایی که در لایه اول جریان غیر صفر (برابر با یک) دارند:

$(s. A_0).\ (s. A_2). (s. A_5). (s. A_6)$

کمان‌هایی که در لایه دوم جریان غیرصفر (برابر با یک) دارند:

$(A_0. B_{02}). (A_0. B_{05}). (A_0. B_{06}).$
$(A_2. B_{02}). (A_2. B_{25}). (A_2. B_{26}).$
$(A_5. B_{05}). (A_5. B_{25}). (A_5. B_{56}).$
$(A_6. B_{06}). (A_6. B_{26}). (A_6. B_{56})$

کمان‌هایی که در لایه سوم جریان غیرصفر (برابر با یک) دارند:

$(B_{02}. C_2). (B_{05}. C_5). (B_{06}. C_6).$
$(B_{25}. C_3), (B_{26}. C_4). (B_{56}. C_1)$

دقت کنید که جواب PDP و همچنین شبکه جریان متناظر منحصر به فرد نیست. در مثال مذکور

$X = \{x_0 = 0. x_1 = 2. x_2 = 7. x_3 = 10\}$

جواب دیگری برای PDP محسوب می‌شود و جریان شدنی متناظر با آن به صورت زیر است:

کمان‌هایی که در لایه اول جریان غیرصفر (برابر با یک) دارند:

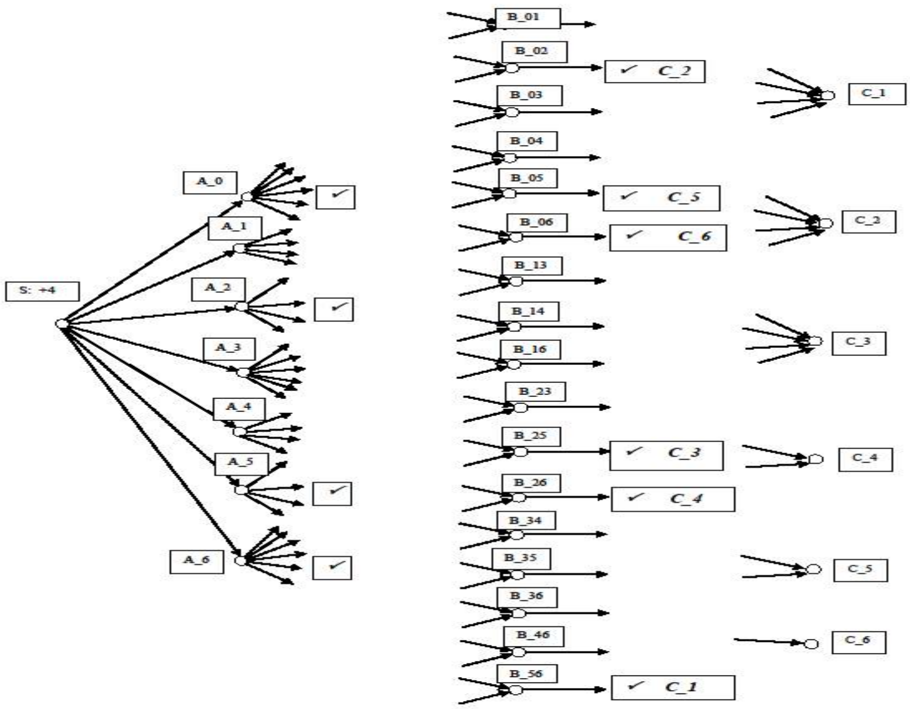
$(s. A_0). (s. A_1). (s. A_4). (s. A_6)$

کمان‌هایی که در لایه دوم جریان غیرصفر (برابر با یک) دارند:

$(A_0. B_{01}).\ (A_0. B_{04}).\ (A_0. B_{06}).$
$(A_1. B_{01}).\ (A_1. B_{14}).\ (A_1. B_{16}).$
$(A_4. B_{04}). (A_4. B_{14}).\ (A_4. B_{46}).$
$(A_6. B_{06}).\ (A_6. B_{16}). (A_6. B_{46})$

کمان‌هایی که در لایه سوم جریان غیرصفر (برابر با یک) دارند:

$(B_{01}. C_1).\ (B_{04}. C_4). (B_{06}. C_6).$
$(B_{14}. C_3). (B_{16}. C_5).\ (B_{46}. C_2)$



شکل ۴: مدل شبکه جریان و یک جریان شدنی متناظر یک جواب مساله PDP. گره های فعال در لایه دوم و سوم که با علامت تیک مشخص شده‌اند، تمامی یالهای ورودی و خروجی‌شان اشباع شده‌اند. نام C_j که جلوی گره‌های انتخاب شده از لایه سوم آمده‌اند، گره مقصد را نشان می‌دهند.

نتیجه‌گیری

در این مقاله یک مدل جدید با استفاده از شبکه جریان برای مسأله PDP ارائه کردیم. در این شبکه، جریان‌های شدنی را پیدا کردیم و نشان دادیم که هر یک از این جریان‌ها جوابی برای مسأله‌ی PDP خواهد بود. برای حل مساله آن را به یک مساله حداکثر جریان صفر و یک تعمیم یافته تبدیل کردیم.

روش‌های کلاسیک مساله حداکثر جریان در زمان چند جمله‌ای حل می‌شوند و با توجه به اینکه تاکنون هیچ الگوریتم چند جمله‌ای برای مساله PDP ارایه نشده است، عدم توفیق این الگوریتم ها دور از ذهن نخواهد بود. برای مساله حداکثر جریان صفر و یک نیز روش‌های چندجمله‌ای وجود دارد[1]. لیکن برای مسایل شبکه جریان صفر و یک تعمیم یافته از الگوریتم‌های برنامه‌ریزی صفر و یک استفاده می‌شود و برتری مدل‌های شبکه جریان در به کار گیری این الگوریتم‌ها، ماتریس ضرایب تکنولوژی ساده آنهاست.

مدل‌های ریاضی در مسائل بهینه سازی دارای پیشینه‌ی زیادی هستند و روش‌های قدرتمندی نیز برای حل آنها وجود دارد. با توجه به اینکه کلاس پیچیدگی مسأله PDP هنوز هم یک مسأله‌ی باز محسوب می‌شود، امیدواریم این مدل نکات و دیدگاه‌های خوبی برای مسأله پیش روی پژوهشگران قرار دهد. لازم به تاکید است روش ارایه شده در این مقاله کلاس پیچیدگی مسأله‌یPDP را مشخص نمی‌کند و کلاس پیچیدگی این مساله به عنوان یکی از مهمترین و قدیمی‌ترین مسایل باز حوزه پیچیدگی مسایل باقی می‌ماند.

فهرست منابع

[1] R. K. Ahuja; T. L. Magnanti.; J. B. Orlin, (1993). Network flow: theory, algorithms, and applications. Prentice Hall.

[2] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. (1993). Nonlinear programming theory and algorithms, Second edition, John Wiley and Sons, Inc. p:512.

[3] M. Cieliebak, S. Eidenbenz, and P. Penna. (2005). Partial Digest is hard to solve for erroneous input data. Theory. Computer. Sci., 349(3): 361-381.

[4] T. Dakic, (2000). On the Turnpike Problem. PhD thesis, Simon Fraser University.

[5] P. Lemke, M. Werman. (1988). on the complexity of inverting the autocorrelation function of a finite integer sequence, and the problem of locating n points on a line, given the $\binom{n}{2}$ unlabeled distances between them. Preprint 453, Institute for Mathematics and its Application IMA.

[6] A. L. Patterson. (1935). A direct method for the determination of the components of interatomic distances in crystals, Zeitschr. Krist. 90: 517-542,

[7] S. S. Skiena, W. Smith, and P. Lemke. (1990). Reconstructing sets from interpoint distances. In Proc. of the 6th ACM Symposium on Computational Geometry (SoCG 1990), pages 332-339.

[8] Z. Zhang. (1994). An exponential example for a partial digest mapping algorithm. Journal of Computational Biology, 1(3):235-239.

[9] Salehi Fathabadi, H. Nadimi, R. (2010). A Continuous Optimization Model for Partial Digest Problem. Journal of Sciences, Islamic Republic of Iran 21.2

[10] Nadimi, Reza, Hassan Salehi Fathabadi, and Mohammad Ganjtabesh. (2011). A fast algorithm for the partial digest problem. Japan journal of industrial and applied mathematics 315-325.