

استفاده از رگرسیون خطی در حل مدل طراحی شبکه دوسطحی

*محمود صدرا ۱ و مهدی زعفرانی ۲

۱دانشجوی ارشد علوم کامپیوتر، دانشگاه حکیم سبزواری، سبزوار، ایران

۲گروه ریاضی کاربردی، دانشگاه حکیم سبزواری، سبزوار، ایران

چکیده

ز جمله راه های کنترل ترافیک شهری، احداث خیابانهای جدید می باشد. اما با توجه به محدودیت همیشگی بودجه، همواره پاسخ این سوال که از میان چندین خیابان کاندیدا برای احداث، کدام موارد انتخاب شود تا بیشترین اثر را در کاهش بار ترافیکی معابر داشته باشد، بسیار حیاتی است. این مساله، مساله "طراحی شبکه حمل و نقل گسسته DNDP نامیده می شود. در این مقاله یک مساله دو سطحی معرفی می شود که هدف آن انتخاب بهینه یالهای جدید برای احداث در یک شبکه حمل و نقل است. از یک روش ترکیبی یادگیری ماشین نظارت شده برای حل این مساله دوسطحی استفاده شده است. مثالهای عددی در محیط پایتون پیاده سازی شده است. برای آماده سازی شبکه شهری از نرم افزار QGIS و پلاگین AequilibraE آن بهره برده ایم.

واژه های کلیدی:

ساله طراحی شبکه حمل و نقل گسسته (DNDP)، برنامه ریزی دوسطحی، مساله تخصیص ترافیک (TAP)، رگرسیون خطی، یادگیری ماشین نظارت شده

۱ مقدمه

مسائل طراحی شبکه حمل و نقل (TNDP) به سه گروه عمده تقسیم می شوند که عبارتند از مسائل طراحی شبکه حمل و نقل گسسته (DNDP) که با پارامترهای گسسته شبکه همچون اضافه کردن یا کم کردن یک یال مرتبط هستند، مسائل طراحی شبکه حمل و نقل پیوسته (CNDP) که با پارامترهای پیوسته شبکه همچون افزایش یا کاهش ظرفیت یک یال مرتبط هستند، مسائل طراحی شبکه حمل و نقل مخلوط (MNDP) که بطور همزمان به هر دو نوع ویژگی می پردازند.

یکی از مسائل معروف DNDP مساله انتخاب یک سبب بهینه از میان چندین خیابان کاندیدا برای احداث است تا بیشترین اثر را در کاهش بار ترافیکی معابر داشته باشد. مساله DNDP یک مساله دو سطحی است که در آن ابتدا مدیران شهری اقدام به انتخاب خیابانهای جدید برای احداث می کنند و سپس مسافران با توجه به شرایط جدید و براساس اصول دوگانه وردراپ و مساله تخصیص ترافیک (TAP) اقدام به انتخاب مسیرها نموده و جریان ترافیک معابر را شکل می دهند. در مساله سطح بالا هدف بیشترین کاهش زمان کل سفر شبکه و هزینه لازم برای احداث خیابانهای جدید است.

۲ مساله تخصیص ترافیک

طراحی یک سیستم حمل و نقل شهری در چهار مرحله مطالعه و انجام می شود که شامل این مراحل است تولید سفر، توزیع سفر، تنوع سفر، تخصیص سفر [4].

با توجه به تقاضاهای سفری که در مرحله تولید سفر بدست آورده ایم، میزان سفر بین تمام نقاط را محاسبه می کنیم. به اینصورت که تمام نقاط را بصورت زوجهای مبدا-مقصد (OD) در نظر می گیریم و حجم سفر مورد تقاضا میان آنها را مشخص می کنیم. منظور از نوع سفر، وسیله نقلیه ای است که مسافر از آن استفاده می کند.

با استفاده از اطلاعات سیستم حمل و نقل شهری، میزان ترافیکی که از هر خیابان و جاده عبور می کند را مشخص می کنیم و آن را "مساله تخصیص ترافیک" می نامیم. این کار بر اساس اصول دوگانه وردراپ انجام می شود. این دو اصل را وردراپ در سال ۱۹۵۲ معرفی کرد که بر اساس آنها سیستم در حالت پایدار هم از دید مسافران و هم از دید مدیر سیستم بهینه عمل می کند [4].

اصل اول: در حالتی که سیستم پایدار شده است همواره مسافران از مسیر که کمترین هزینه را دارد سفر خود را انجام می دهند و هیچ مسافری نمی تواند با انتخاب مسیر دیگری هزینه سفرش را کمتر کند.

اصل دوم: در حالتی که سیستم پایدار شده است متوسط هزینه سفر در کل سیستم کمترین مقدار می باشد

۱-۲ تعریف مساله تخصیص ترافیک :

مساله تخصیص ترافیک بصورت مساله بهینه سازی محدب زیر فرمول نویسی می شود (نمادهای مورد استفاده در ضمیمه مقاله آمده است):

$$\min \sum_{a \in A} \int_0^{x_a} s_a(t) dt,$$

subject to

$$\sum_{p \in P_k} h_{pk} = r_k, \forall k$$

$$x_a = \sum_k \sum_{p \in P_k} \delta_{kap} h_{pk}, \forall a \in A$$

$$h_{pk} \geq 0, \forall a \in A, \forall k$$

ما در این مقاله برای حل مدل ریاضی مساله از بسته pyomo و حل کننده های ipopt و glpk در پایتون استفاده و به کمک آن مساله تخصیص ترافیک را بعنوان یک مساله بهینه سازی ریاضی حل کردیم. برای آشنایی با روشهای حل Tap می توانید به [4] نگاه کنید.

۳ مسائل برنامه ریزی دو سطحی

مساله بهینه سازی دو سطحی در واقع مربوط به برنامه اقتصادی استکلبرگ در نظریه بازیهاست. بازیکن اول که تابع هدف سطح اول (پیشرو) را دارد، سعی می کند با انتخاب مقدار برای متغیر تصمیم y ، مقدار تابع $F(x,y)$ را کمینه کند، بازیکن دوم که تابع هدف سطح دوم (پیرو) را دارد با اطلاع کامل از تصمیم بازیکن اول سعی می کند با انتخاب مقدار برای متغیر تصمیم x ، تابع هدف $f(x,y)$ را کمینه کند. به این ترتیب تصمیم بازیکن دوم وابسته است به تصمیم بازیکن [1]. ساختار کلی برنامه ریزی دوسطحی بصورت زیر می باشد:

$$\min_{x \in X, y \in Y} F(x, y)$$

$$s.t \quad G(x, y) \leq 0$$

$$\min_{y \in Y} f(x, y)$$

$$s.t \quad g(x, y) \leq 0$$

برخی از روشهای حل مسائل دو سطحی عبارتند از: روش نقطه راسی، روش شاخه و کران، روش محور گیری مکمل، روش تابع جریمه، روش ناحیه ایمن، روش رویکرد کاهشی، روش KKT، یادگیری ماشین [3].

۴ مساله طراحی شبکه حمل و نقل گسسته (DNDP)

مساله NDP به روش مشخص کردن تعدادی از یالهای بهینه بر روی یک شبکه اطلاق می شود. در این مقاله یک مساله DNDP دو سطحی بررسی می شود. در سطح اول دیدگاه مدیران شهری جهت انتخاب یک سبب بهینه از خیابانهای (یالها) جدید مورد بررسی قرار می گیرد. در سطح دوم دیدگاه کاربران شبکه جهت انتخاب بهترین مسیرهای ممکن بین مبدا و مقصدهای متفاوت در قالب مساله

TAP بررسی می شود. در مدل کلاسیک DNDP هدف انتخاب یک سید بهینه از خیابانهای جدید با توجه به یک بودجه محدود است. در این مقاله محدودیت بودجه بعنوان قسمتی از تابع هدف سطح اول در نظر گرفته شده است.

$$\min_{x_a, y_a} Z(x_a, y_a) = \sum_{a \in A \cup A'} x_a \cdot t_a(x_a) + \sum_{a \in A'} c_a \cdot y_a$$

subject to

$$y_a = \{0, 1\}, \quad a \in A'$$

$$\min_x \sum_{a \in A \cup A'} \int_0^{x_a} t_a(x_a) dx,$$

subject to

$$\sum_{k \in P_i} h_k = q_i, \quad i \in I$$

$$x_a = \sum_{i \in I} \sum_{k \in P_i} h_k \delta_{a,p}, \quad a \in A \cup A'$$

$$x_a \geq 0, \quad a \in A \cup A'$$

$$x_a \leq M \cdot y_a, \quad a \in A'$$

تابع هدف عبارت از مجموع زمان سفر کل شبکه (T1) و هزینه کل احداث یالهای انتخابی (T2) است. برای متناسب سازی تاثیر دو جمله تابع هدف، از شکل نرمال شده آنها استفاده کرده ایم.

$$T1 = \sum_{a \in A \cup A'} x_a \cdot t_a(x_a)$$

$$T2 = \sum_{a \in A'} c_a \cdot y_a$$

$$\min_{x_a, y_a} Z(x_a, y_a) = \frac{\sum_{a \in A \cup A'} x_a \cdot t_a(x_a) - \min(T1)}{\max(T1) - \min(T1)} + \frac{\sum_{a \in A'} c_a \cdot y_a - \min(T2)}{\max(T2) - \min(T2)}$$

۵ روش حل

ما در حل مساله از رگرسیون خطی که یکی از کاربردهای مهم جبر خطی در هوش مصنوعی و یادگیری ماشین است استفاده کرده ایم. رگرسیون خطی نوعی تابع پیش بینی کننده خطی است که در آن متغیر وابسته (متغیری که قرار است پیش بینی شود) به صورت ترکیبی خطی از متغیرهای مستقل پیش بینی می شود. بدین معنی که هر کدام از متغیرهای مستقل در ضربی که در فرایند تخمین برای

آن متغیر به دست آمده ضرب می‌شود و جواب نهائی مجموع حاصل ضرب‌ها به علاوه یک مقدار ثابت خواهد بود که آن هم در فرایند تخمین به دست آمده است.

ساده‌ترین نوع رگرسیون خطی، رگرسیون خطی ساده است که بر خلاف رگرسیون خطی چندگانه، تنها یک متغیر مستقل دارد. نوع دیگر رگرسیون خطی، رگرسیون خطی چندمتغیره است که در آن به جای پیش‌بینی یک متغیر وابسته چندین متغیر وابسته پیش‌بینی می‌شود. ما از رگرسیون خطی چندگانه استفاده کرده ایم و برای این منظور تمام داده ها را بصورت بردار و ماتریس در آورده ایم تا بتوانیم از الگوریتمهای جبر خطی برای پیش بینی مقادیر متغیرهای وابسته استفاده کنیم. متغیر وابسته مقدار تابع هدف است و متغیرهای مستقل همان یالهای کانیدیا می باشند.

ابتدا مقادیر کمینه و بیشنه هر کدام از جملات تابع هدف به تنهایی را محاسبه می کنیم که مربوط می شود به حالاتی که تمام یالها انتخاب شوند یا هیچکدام از یالها انتخاب نشوند. سپس به سراغ یک فرایند تکراری محاسبه مقدار تابع هدف با توجه به یالها انتخاب شده می رویم به اینصورت که ابتدا یک بردار شدنی از یالهای کانیدیا (Y) در نظر می گیریم که ما بردار Y تمام صفر را در نظر گرفته ایم و با استفاده از آن تابع سطح پایین را که TAP است حل می کنیم. سپس جریانهای بدست آمده را به تابع سطح بالا می فرستیم تا مقدار تابع هدف (Z) را بدست آید. حالا یک مجموعه از یالها داریم و به ازای آنها یک مقدار برای تابع هدف. حال به کمک این دو مقدار یک تابع رگرسیون می سازیم. و از روی آن، مقدار جدیدی برای Y حدس می زنیم و آنرا به تابع سطح پایین TAP می فرستیم تا جریانهای جدید را محاسبه کند و به تابع بالا بفرستد و مقدار جدیدی برای تابع هدف بدست آورد حالا به کمک این دو مقدار و مقادیر قبلی Y و Z یک تابع رگرسیون جدید می سازیم و دوباره مقدار جدیدی برای Y حدس می زنیم و کار در قالب یک حلقه ادامه پیدا می دهیم.

شرط خروج از حلقه را هم می توان کم شدن اختلاف میان دو Z متوالی از یک حد خاص در نظر گرفت و هم تعداد تکرار. در مثالهای عملی شرط تعداد تکرار بهتر و موثرتر عمل کرده است. نمودار روش در ضمیمه مقاله آمده است.

۶ نتایج اصلی

در این مقاله یک مدل دو سطحی طراحی شبکه مورد بررسی قرار گرفته است که یک مدل np -hard است و حتی بدست آوردن یک جواب بهینه محلی برای آن نیز np -hard است. ما به کمک یک روش یادگیری ماشین نظارت شده، این مساله را حل کردیم و بر روی یک شبکه نسبتاً بزرگ جوابهای مناسب در زمانها بسیار کوتاه بدست آوردیم.

۷ نتایج عددی

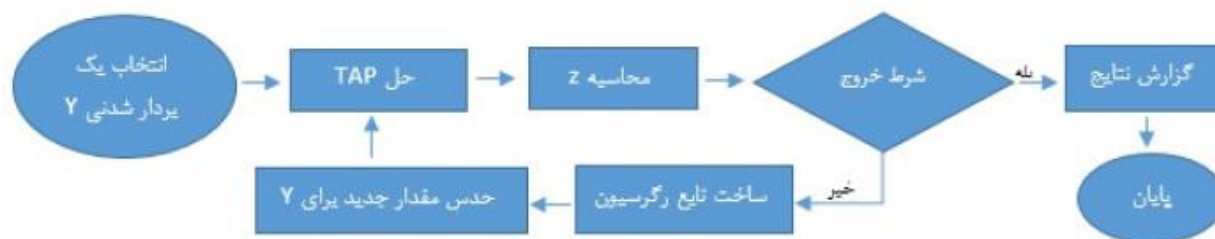
برای مطالعه موردی شبکه شهر Sioux Falls که یک الگوی استاندارد و شناخته شده در حوزه حمل و نقل است، استفاده شده است که در این آدرس (<https://github.com/bstabler/TransportationNetworks>) در دسترس است. نقشه شهر Sioux Falls که با نرم افزار QGIS تهیه شده است و در آن تمام یالها یک طرفه هستند در ضمیمه مقاله آمده است. مساله این است که ۱۰ یال با هزینه های مشخص، کانیدای احداث هستند. هدف یافتن سببی از یالها است که با کمترین هزینه، بیشترین تاثیر را در کاهش ترافیک داشته باشند. جدول یالهای کانیدیا و هزینه آنها در ضمیمه آمده است. ما از یک سیستم با ۶ گیگا بایت حافظه و پردازنده Core i5 نسل ۵ برای اجرای برنامه استفاده کرده ایم و پس از ورود اطلاعات، برنامه را اجرا کردیم که در یک حلقه با ۲۰ تکرار مساله را حل کرد. مدت زمان اجرا ۲۳ دقیقه و مقدار بیشینه تابع هدف ۵۹۷۳/۲۰۱ و مقدار کمینه تابع هدف ۰/۴۳۷۷۴۸۷۳ گزارش شد که در تکرار ۱۷ام اتفاق افتاد.

جدول ۱: هزینه احداث یالها.

شماره یال	هزینه احداث (میلیون دلار)
۱	۴۴
۳	۵۹
۱۲	۷۹
۱۵	۸۶
۳۳	۵۸
۳۶	۴۴
۴۱	۵۹
۴۶	۷۹
۵۰	۸۶
۵۵	۵۸

$G = (A, N)$	یک شبکه با مجموعه N از گره ها و مجموعه A از یالها
A'	مجموعه یالهای کاندیدا برای ایجاد
$K \subseteq N \times N$	مجموعه زوجهای مبدا-مقصد شبکه G
P_k	مجموعه مسیرهای ساده بین نود ابتدایی و انتهای k امین زوج
h_{pk}	جریان موجود در مسیر p ام در مجموعه P_k
x_a	جریان یال a
$s_a(x_a)$	زمان سفری که کاربر در زمان سفر در یال a با آن روبرو می شود
r_k	تقاضای سفر بین K امین زوج
$\Delta = [\delta_{kap}]$	ماتریس واقع شدن یک یال در یک مسیر که در آن δ_{kap} ، یک است اگر مسیر p ام از k امین زوج شامل یال a باشد و در غیر اینصورت صفر
Y	بردار یالهای کاندیدا برای اضافه شدن به شبکه (1: یال اضافه شود. 0: یال اضافه نشود)
C_a	هزینه احداث یال a
B	کل بودجه
q_i	تقاضای سفر زوج i ام
$g_k(\pi)$	تابع تقاضای سفر روی زوج مبدا-مقصد k ام
M	عدد بسیار بزرگ

شکل ۱: جدول نمادهای بکار رفته در مقاله .



شکل ۲: نمودار روش حل .

یالهای انتخاب شده برای احداث										مقدار تابع هدف	شماره تکرار
1	3	12	15	33	36	41	46	50	55		
*										1.1031999	1
					*					1.0168463	2
									*	1.10262456	3
				*						1.02006345	4
				*	*					0.93859915	5
	*			*	*					1.2005973	6
				*	*				*	1.07822271	7
				*	*	*				0.90122969	8
					*	*				1.01080675	9
				*		*				0.93173014	10
						*				0.93362567	11
				*	*	*	*			0.49705059	12
		*	*	*	*	*	*	*		0.48706806	13
		*	*	*		*	*	*		0.46927982	14
			*	*		*	*	*		0.49038387	15
		*		*		*	*	*		0.49293671	16
		*	*	*		*	*			<u>0.43774873</u>	17
		*	*	*	*	*	*			0.44185862	18
			*	*		*	*			0.47827769	19
		*		*		*	*			0.49926913	20

شکل ۳: نتایج ۲۰ تکرار .

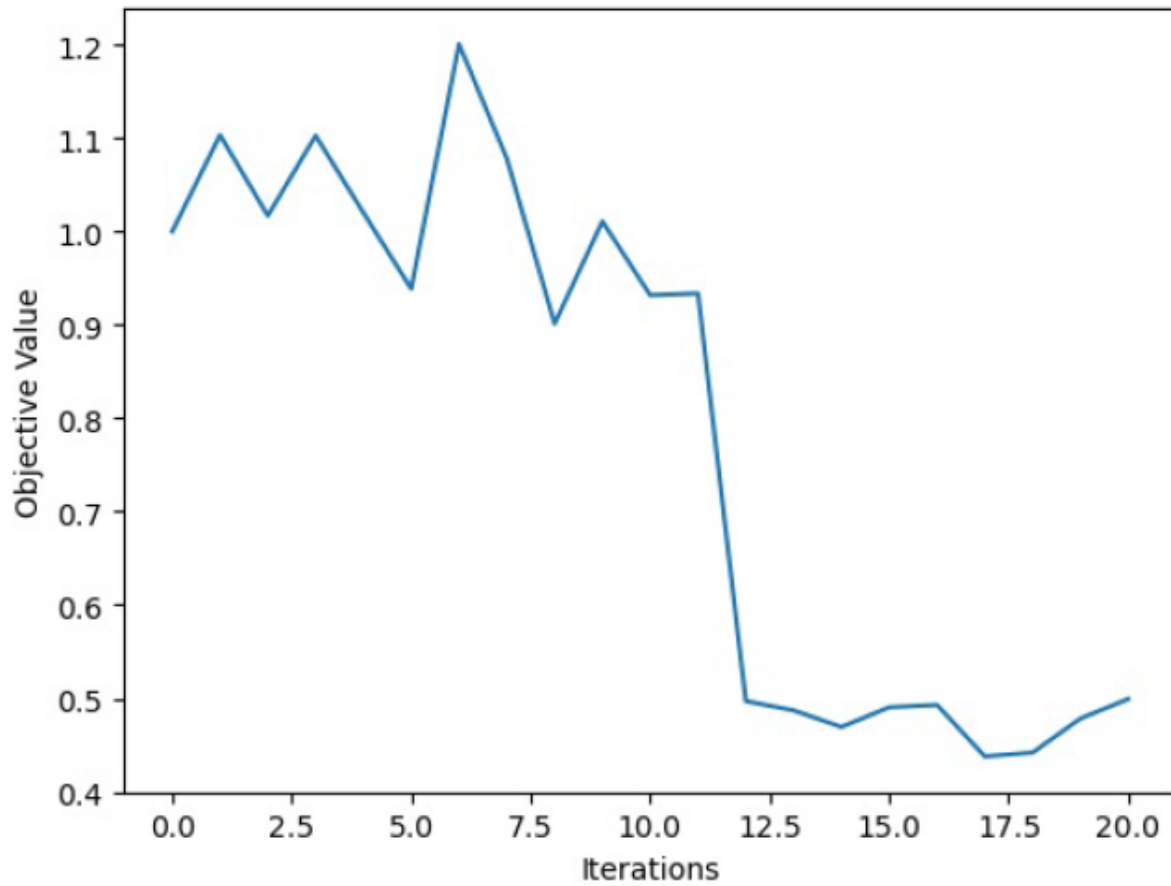
مراجع

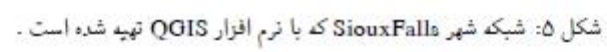
- [1] Abareshi, Maryam, and Mehdi Zaferanieh. "A bi-level capacitated P-median facility location problem with the most likely allocation solution." *Transportation Research Part B: Methodological* 123 (2019): 1-20.
- [2] Bagloee, Saeed Asadi, et al. "A hybrid machine-learning and optimization method to solve bi-level problems."

Expert Systems with Applications 95 (2018): .142-152

[3] Dempe, Stephan, and Alain Zemkoho. Bilevel optimization. Springer, .2020

[4] Patriksson, Michael. The traffic assignment problem: models and methods. Courier Dover Publications, 2015





شکل ۵: شبکه شهر Sioux Falls که با نرم افزار QGIS تهیه شده است.