فصلنامه علمی مهندسی ترافیک
| JOURNAL OF TRAFFIC ENGINEERING |

ارائه مدل پیش بینی تاخیر پرواز

عبدالرضا شیخ الاسلامی[۱]، حسن خاکسار[۲]، علی فغانی[۳]

۱- دکتری برنامه ریزی حمل و نقل، دانشگاه علم و صنعت ایران

۲- دانشجوی دکتری برنامه ریزی حمل و نقل، دانشگاه علم و صنعت ایران

۳- کارشناس ارشد برنامه ریزی حمل و نقل، دانشگاه علم و صنعت

چکیده

تاخیر در پروازها و تلاش برای کاهش آنها، یکی از مواردی است که همواره مورد توجه محققین و پژوهشگران برنامه ریزی حمل و نقل و تحقیق در عملیات بوده است. به دلیل ماهیت تصادفی و غیر قطعی پارامترهای موثر در زمان بندی برنامه پروازی، شرکتهای هواپیمایی معمولا با پدیده تاخیر مواجه هستند که در بسیاری از موارد، عوامل تولید کننده این تاخیرها و مدیریت آنها در اختیار شرکتهای هواپیمایی نیستند. آنچه موضوع را بحرانی تر می‌کند تاثیرپذیری پروازها از هم و انتشار (رشد) تاخیرها در شبکه پروازی شرکتهای هواپیمایی است. در این مقاله به بررسی عوامل موثر در بروز تاخیرهای پروازها در شرکتهای هواپیمایی پرداخته شده است و در نهایت یک مدل داده کاوی برای پیش بینی تاخیرهای پروازی ارائه شده است. مدل پیشنهادی بر روی داده‌های پروازی یکی از شرکتهای هواپیمایی اجرا شده است. بر اساس نتایج مدلسازی، مدل پیشنهادی قابلیت پیش بینی وقوع تاخیر و میزان آن در ۷۰٪ مواقع را دارد.

کلید واژه: تاخیر پرواز، مدل استوار، برنامه ریزی پرواز، تاخیر انتشاری (عکس العملی)

۱- مقدمه

در دهه‌های اخیر مسئله برنامه ریزی پرواز مورد توجه محققین برنامه ریزی حمل و نقل، ریاضیات و تحقیق در عملیات بوده است. دو مانع چالشهای اساسی در توسعه ابزار برنامه ریزی قابل رهگیری در صنعت حمل و نقل هوایی ایجاد کرده‌اند. اولی ابعاد مسئله و پیچیدگیهای شبکه پروازی است. مثلاً یک شرکت هواپیمایی بزرگ ممکن است حدود ۵۰۰ هواپیما و بیش از هزار پرواز روزانه در شبکه‌ای با دهها هزار جفت مبداء- مقصد داشته باشد و میلیونها مسافر در هر سال از این شبکه استفاده کنند. تعداد زیاد تصمیمهایی که باید گرفته شود و تعاملات پیچیده بین منابع مختلف (خدمه، هواپیما و فرودگاه و ...) فرآیند برنامه ریزی را تبدیل به یک وظیفه چالش برانگیز می‌کند.

مانع دوم طبیعت پویا و تصادفی این سیستم است. بسیاری از بخشهای سیستم محدود به اختلالهای کلی یا جزئی هستند. مثلاً زمان یک پرواز به سرعت روزانه باد وابسته است. زمان خزش آن به تراکم فرودگاه وایسته است و موارد تعمیر و نگهداری و سایر دلایل ممکن است برخاست را با تاخیر مواجه کند. از آنجا که مدلهای حمل و نقل هوایی قصد دارند بهره‌وری از منابع را بیشینه کنند و ممکن است تغییرپذیری در نظر گرفته نشود. در این موارد، پلان بهینه ممکن است در صورت بروز اختلال بسیار ضعیف عمل کند چرا که تاخیرها نمی‌توانند جذب شوند و در سیستم منتشر می‌شوند.

لذا و با توجه به رشد فزاینده تقاضای حمل و نقل هوایی و ثابت ماندن فضای عرضه، لزوم برنامه ریزی و بهینه سازی این صنعت پیش از پیش احساس می‌شود [۱]. پیشرفتهای اخیر در صنعت رایانه و توسعه الگوریتمهای فرا ابتکاری موجب شده است مسئله برنامه ریزی پرواز در سالهای اخیر مورد توجه بیشتری قرار گیرد. مدلهای column generation و branch and bound یکی از روشهایی است که برای حل شبکه‌های بزرگ مقیاس نظیر شبکه‌های حمل و نقل هوایی مورد استفاده قرار می‌گیرند [۲]. برنامه ریزی پرواز شامل دو فاز برنامه ریزی و عملکردی می‌شود [۳]. در فاز برنامه ریزی چهار مرحله کلی انجام می‌شود:

-زمان بندی پرواز

-تخصیص ناوگان / مسیریابی

-تعمیر و نگهداری

-تخصیص خدمه

در فاز عملکردی سه بخش اصلی وجود دارد:

-مدیریت درآمد

-تخصیص گیت پرواز

-عملیات در شرایط اختلال

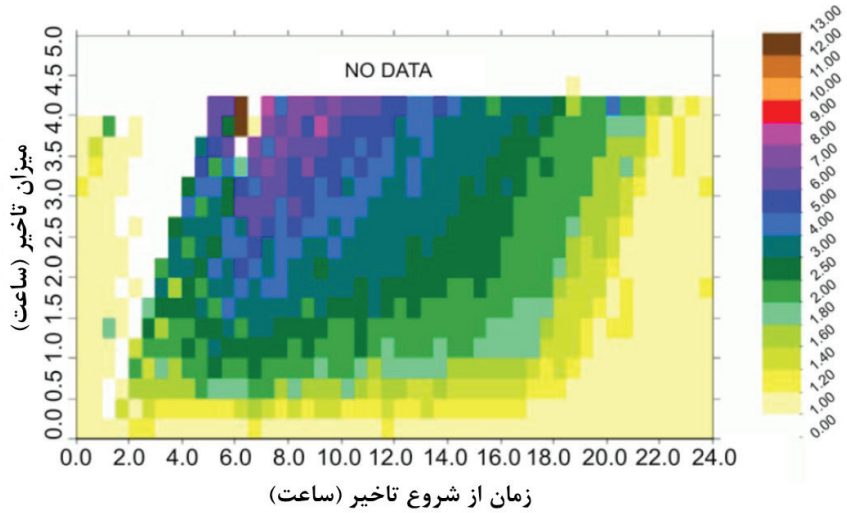
مثالهایی از ادبیات موضوع برنامه ریزی یکپارچه شامل مطالعات کوردئو و همکاران [۴]، کوهن و بارنهارت، [۵] و کلابجان، و همکاران [۶] می‌شود. لوهاتپونانت، و بارنهارت [۷]، رگسینگ، و همکاران [۸]، استویکوویچ، و سومیس، [۹] و سندهو و کلابجان [۱۰] هر کدام بخشهایی از فرآیند برنامه ریزی پرواز را به صورت یکپارچه حل کرده‌اند. مرور بیشتری از ادبیات موضوع مدلهای بازیابی در مقاله‌های براتو و بارنهارت و کوهل و همکاران آورده شده است [۱۱] و [۱۲].

۲– تعریف مسئله

یکی از گامهای اصلی در برنامه ریزی استوار پرواز پیش بینی تاخیر در شبکه پروازی است. برای افزایش انعطاف پذیری شبکه پروازی در مقابل اختلالها لازم است وقوع تاخیر و میزان این تاخیرها در شبکه شناسایی شوند. شناسایی تاخیرها در شبکه قابلیت برنامه ریزی برای مقابله با تاخیرها از طریق در نظر گرفتن راهکارهایی جهت جذب تاخیرها را ممکن خواهد ساخت. بنابراین اولین گام در حل مسئله برنام ریزی پرواز، پیش بینی تاخیرها است. لغو یک پرواز می‌تواند اثرات مخربی روی عملیات برنامه ریزی شده بگذارد و مواردی مانند انفصال در ارتباطات پروازی مسافران و خدمه و ایجاد مشکلات روی شبکه را بوجود آورد. مشکلات جوی نیز سطح ایجاد گسست در شبکه پروازی را بوسیله دو عامل زیر افزایش می‌دهد:

۱-برنامه زمانی خطوط تحت شرایط عادی آب و هوایی و انجام عملیاتی بدون هر گونه توقف زمانی در برنامه، جهت بیشینه کردن بازده منابع، بهینه شده است. بنابراین نیازمند فراهم کردن شرایطی برای جذب و از بین بردن اثرات تاخیرات می‌باشد.

۲-رشد بی رویه ترافیک هوایی در شرایطی رخ می‌دهد که ظرفیت تعدادی از فرودگاههای پر تراکم، ثابت مانده است. هزینه‌های عملیاتی مستقیم پروازها در سال ۲۰۰۷ در حدود ۸ میلیارد دلار تخمین زده شده است در حالی که هزینه‌های اضافی ناشی از تاخیر مسافران حدود ۴ میلیارد دلار برای هر سال برآورد شده است. به دلیل مشکلات اقتصادی سالهای اخیر خطوط هوایی و افزایش هزینه سوخت، مدیریت این مشکلات



شکل ۱: برآورد انتشار تاخیرها در طول زمان

در کنترل هزینه‌های پیش بینی نشده آنها، موثر است.

تاخیر پرواز ممکن است به دلیل اتصال منابع به یکدیگر در کل شبکه انتشار یابد. برای مثال، هنگامی که پرواز مشخصی می‌خواهد مسافران، کالا و خدمه خود را دریافت نماید ممکن است از تاخیر ۴۰ دقیقه‌ای پرواز متصل به آن تاثیر بپذیرد. در شکل (۱) نحوه انتشار تاخیرها در طول زمان در شبکه خطوط هوایی آورده شده است. بر این اساس بیشترین زمان انتشار تاخیر معمولاً بین ۶ تا ۸ ساعت اولیه پس از رخ دادن تاخیر است و بعد از ۱۰ ساعت انتشار تاخیر شروع به استهلاک و از بین رفتن می‌کند [۱۳].

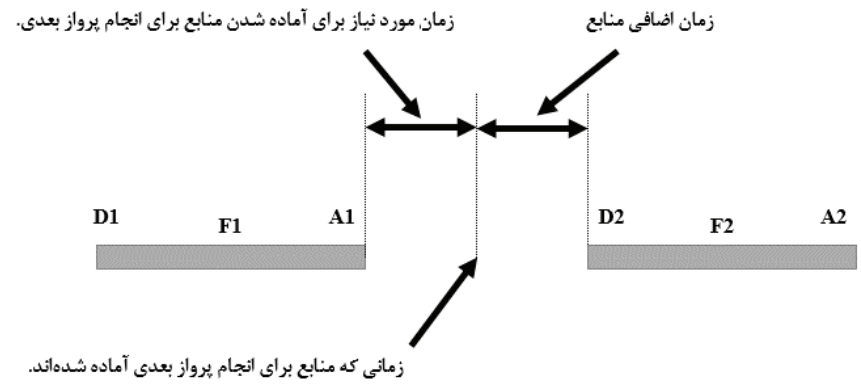
منابع اصلی شرکتهای هواپیمایی در برنامه ریزی پرواز عبارتند از: هواپیما و خدمه. معمولاً این منابع به گونه‌ای زمان بندی می‌شوند که به مجموعه متوالی از پروازها خدمات دهند. یک تاخیر در یک منبع از یک پرواز منجر به تاخیر در یک یا چند پرواز متوالی بعد از آن می‌شود. پس از رسیدن یک هواپیما یا خدمه، آن هواپیما یا خدمه مجاز است طی زمان مشخصی برای انجام پرواز بعدی آماده شود (حداقل زمان چرخش مجاز). مثلاً در مورد یک هواپیما این زمان بندی برای پیاده شدن مسافران، پیاده‌کردن بارها و چمدانها، تمیزکردن هواپیما، سوخت‌گیری، بارگیری چمدانهای پرواز بعدی، سوار شدن مسافران و موارد مشابه لازم است. مشابهاً برای یک خدمه این زمان مورد نیاز است تا از گیت پرواز ورودی به گیت پرواز بعدی پیاده روی کنند. اگر منابع برای پرواز بعدی، قبل از زمان بندی حرکت آن حاضر و آماده باشد زمان اضافی منبع موجود است. زمان اضافی این منبع تفاوت بین زمان حرکت پرواز خروجی و آماده شدن منابع آن تعریف شده است. در شکل ۲ مثالی از یک زمان اضافی بین اتصال یک منابع بین پروازهای F۱ و F۲ ارائه شده است که با دو مستطیل نشان داده شده است. پس از اینکه منابع به پرواز F۱ در A۱ رسیدند یک بازه زمانی از پیش تعریف شده زمانی به این منبع داده شده است تا برای پرواز بعدی آماده شوند. زمان اضافی تفاوت بین زمان حرکت F۲ و D۲ و زمان آماده شدن این منبع است.

برای تخصیص بهینه زمان اضافی به منابع باید وقوع و میزان تاخیرها پیش بینی شده باشد. تا زمانی که از صحت پیش بینی تاخیرها اطمینان وجود نداشته باشد نمی توان برنامه ریزی استوار را به صورت بهینه انجام داد. بنابراین پیش بینی تاخیر در پروازها از دو جنبه اهمیت دارد؛

- تخصیص زمان اضافی به منابع در برنامه ریزی استوار
- انجام عملیات مدیریت تاخیر

۳- کاربرد داده کاوی در پیش بینی تاخیر

گسترش سیستمهای پایگاهی و حجم بالای داده‌های ذخیره شده در این سیستمها نیاز به ابزاری جهت پردازش و تحلیل و ارائه خروجیهای مورد نیاز از داده‌ها دارد. ابزار معمول گزارش گیری می‌توانند اطلاعات معمول مورد نیاز کاربران را تامین کنند تا بتوانند به نتیجه گیری در مورد داده‌ها و روابط منطقی میان آنها بپردازند. اما با افزایش حجم داده‌ها شناسایی الگو و رابطه منطقی بین داده‌ها عملی نخواهد بود. داده کاوی یکی از مهم ترین و کاربردی ترین روشهایی است که به وسیله آن می‌توان الگوهای مفید در داده‌ها را با حداقل دخالت سلیقه استخراج کرد.



شکل ۲: مثالی از زمان اضافی بین منابع

داده کاوی روشهای گوناگون و متنوعی دارد. مهم ترین روشهای مورد استفاده در پیش بینی تاخیر پروازی در مطالعات قبلی عبارتند از: کلاس بندی و خوشه بندی. هر کدام از این گروهها هم خود دارای روشهای متنوع و متعددی است که مهم ترین موارد و مفاهیم در ارتباط با آنها در ادامه ذکر می‌شود.

کلاس بندی به روشهای زیر امکان پذیر است؛

-طبقه بندی بیز

-درخت تصمیم

-نزدیکترین همسایگی k

-الگوریتمهای ژنتیک

-شبکه‌های عصبی

در این مقاله برای پیش بینی تاخیرها از روش کلاس بندی بیزین استفاده شده است. کلاس بندی بیزین بر مبنای تئوری بیز می‌باشد. روش کلاس بندی بیزین یکی از روشهای دقیق و سریع برای استفاده در داده‌های بزرگ مقیاس است [۱۴]. در این قسمت نیاز است تئوری بیز برای استفاده در کلاس بندی داده‌ها تعریف شود. فرض کنید قرار است بهترین مقدار فرض j در میان مفروضات X پیدا شود. برای این کار باید محتمل ترین با مقدار داده شده X تعیین شود. از $P(j)$ برای معرفی احتمال اولیه درستی فرضیه j قبل از یافتن دادهای آموزشی استفاده می‌شود. $P(j)$ معمولاً احتمال پیشین نامیده می‌شود و بیانگر هر دانش پیشینی است که در مورد درستی فرض j سخن می‌گوید. یعنی اگر هیچ دانش اولیه‌ای از مفروضات j وجود نداشته باشد یک احتمال یکسان به کل فضای مفروضات j تخصیص می‌یابد. $p(X|j)$ بیانگر احتمال X در شرایطی است که فرض j صادق است. در کلاس بندی بیزین به دنبال یافتن $p(X|j)$ هستیم که احتمال پسین نام دارد. قاعده و قانون تئوری بیزین در رابطه (۱) آورده شده است که بیانگر نحوه محاسبه احتمال پسین بر مبنای احتمال پیشین است.

$$p(j|X)=\frac{p(X|j)p(j)}{p(X)} \qquad (۱)$$

در کلاس بندی بیزی پیش بینی تاخیر لازم است مجموعه‌ای از فرضهای J توسط برنامه ریزی در نظر گرفته شود و برنامه ریزی بر مبنای یافتن محتمل ترین جواب j در فضای J انجام شود. کلاس بندی بیزی در این حالت به حداکثر پسین[1] (MAP) معروف است. در این روش می‌توان فرض MAP را با استفاده از قضیه بیز و برای محاسبه احتمال پسین هر کاندیدا مطابق با رابطه (۲) محاسبه کرد.

$$h_{MAP}=\arg_{j\in J}P(j|X) \qquad (۲)$$
$$=\arg_{j\in J}\max\frac{P(X|j)P(j)}{P(X)}=\arg_{j\in J}P(X|j)P(j)$$

کلاس مربوط به نمونه جدید می‌تواند با استفاده از کلاس بندی بیزین محاسبه شود. بنابراین احتمال رخداد i امین نمونه با شناسایی $a_1, a_2 \dots a_n$ محاسبه خواهد شد.

$$v_{MAP}=\arg_{v_j\in V}\max P(v_j|a_1,a_2,\dots,a_n) \qquad (۳)$$

پس از بازنویسی رابطه بر مبنای قاعده بیزین خواهیم داشت؛

$$v_{MAP}=\arg_{v_j\in V}\max\frac{P(a_1,a_2,\dots,a_n|v_j)P(v_j)}{p(a_1,a_2,\dots,a_n)} \qquad (۴)$$
$$=\arg_{v_j\in V}\max P(a_1,a_2,\dots,a_n|v_j)P(v_j)$$

احتمال وقوع هر یک از رخدادهای $a_1, a_2 \dots a_n$ از طریق ضرب کردن آنها محاسبه خواهد شد.

$$v_{NB}=\arg_{v_j\in V}\max P(v_j)\prod_i p(a_i|v_j) \qquad (۵)$$

در این رابطه v_{NB} خروجی طبقه بندی بیزین است. تعداد ترم‌های محاسبه شده در این روش معادل حاصلضرب تعداد کلاسها در تعداد خروجیها می‌باشد که کمتر از تعداد ترمهای خواهد بود.

۴- مطالعه موردی

برای پیاده سازی نتایج استفاده از مدل پیش بینی بیزین در وقوع تاخیرهای برنامه‌های پروازی از داده‌های پروازی یکی از شرکتهای هواپیمایی بزرگ داخلی استفاده شده است. خصوصیات کلی بانک داده مورد استفاده در جدول ۱ آورده شده است و نحوه نامگذاری داده‌ها در مدلسازی در جدول ۲ ارائه شده است. لازم به ذکر است اطلاعات مربوط به شرایط جوی در روزهای انجام پرواز از سازمان هواشناسی استعلام شده است.

1 MAP (Maximum a Posteriori)

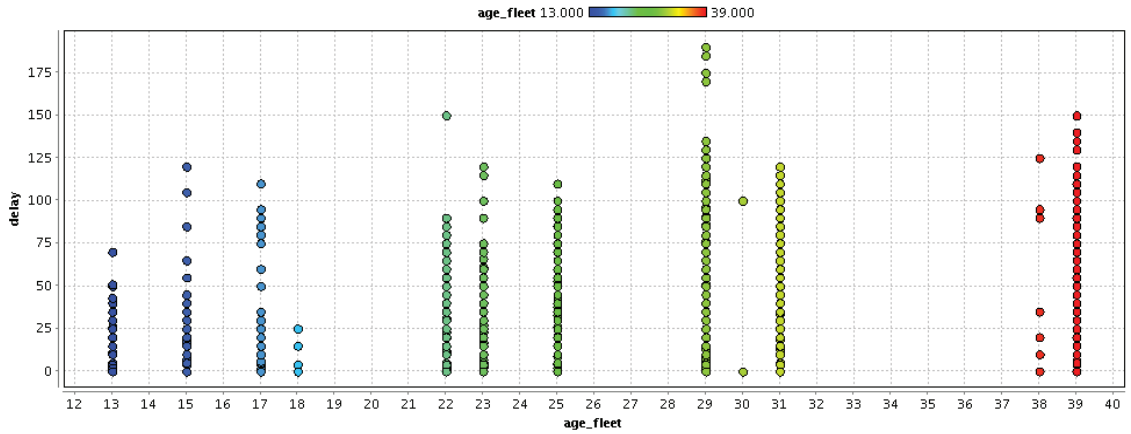
جدول ۱: مشخصات کلی داده‌های استفاده شده در مدلهای پیش بینی تاخیر

عنوان	مشخصات در مطالعه موردی
تعداد فرودگاههای فعال	۵۲
تعداد نوع هواپیما	۹

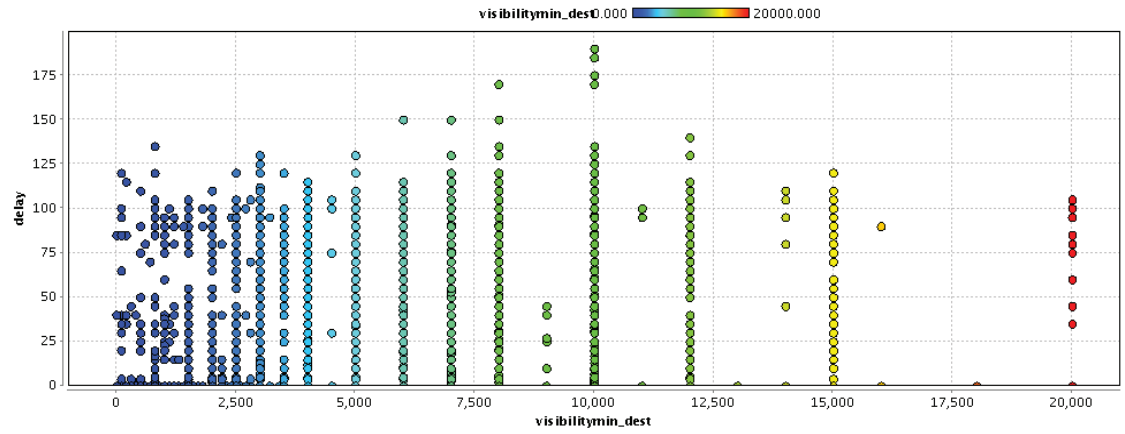
جدول ۲: متغیرهای استفاده شده در مدل پیش بینی تاخیر پروازها

متغیرهای مورد استفاده در داده کاوی	مسافت دید در مبداء و مقصد	visibilitymin_origin, visibilitymin_dest
	زمان پرواز زمان بندی شده	sch_dep
	نوع هواپیما	Act_type
	سن هواپیما	age_fleet
	درجه حرارت مبداء و مقصد	avg _Temp_origin, avg _Temp_dest
	فشار هوای مبداء و مقصد	avg_ Station _pressure_origin, avg_ Station _pressure_dest

در شکل ۳ و ۴ تاثیر دو پارامتر سن ناوگان و مسافت دید بر میزان تاخیرها آورده شده است.



شکل ۳: اثرات سن ناوگان بر میزان تاخیر



شکل ۴: اثرات حداقل مسافت دید در فرودگاه مقصد بر میزان تاخیر

جدول ۳: دقت روش بیزین در پیش بینی وقوع و میزان تاخیر

روش پیش بینی	دقت در پیش بینی بروز تاخیر ٪	دقت در پیش بینی میزان تاخیر ٪
بیزین	۷۰,۱۷	۶۹,۸۱

جدول ۴: ماتریس در هم ریختگی روش بیزین

روش پیش بینی	بازه کلاس (دقیقه)	پارامتر مورد ارزیابی	
		Class precision	Class recall
روش بیزین	۱۵-۰	**۷۶,۳۵**	**۵۱,۶۹**
	۶۰-۱۵	**۶۴,۸۲**	**۵۸,۰۲**
	بیشتر از ۶۰	**۶۹,۷۴**	**۹۴,۰۱**

۵- ارزیابی نتایج

برای ارزیابی کیفیت جوابها از ماتریس درهم ریختگی استفاده شده است که چگونگی عملکرد الگوریتم دسته‌بندی را با توجه به مجموعه داده‌های ورودی به تفکیک انواع دسته‌های مساله دسته‌بندی نشان می‌دهد. در این ماتریس مفاهیم TN، FP، FN و TP به شرح زیر می‌باشند:

TN ؛ تعداد رکوردهایی که دسته واقعی آن‌ها منفی بوده و الگوریتم دسته‌بندی نیز دسته آن‌ها را به درستی منفی تشخیص داده است.

FP ؛ تعداد رکوردهایی که دسته واقعی آن‌ها منفی بوده و الگوریتم دسته‌بندی دسته آن‌ها را به اشتباه مثبت تشخیص داده است.

FN؛ تعداد رکوردهایی که دسته واقعی آن‌ها مثبت بوده و الگوریتم دسته‌بندی دسته آن‌ها را به اشتباه منفی تشخیص داده است.

TP؛ تعداد رکوردهایی که دسته واقعی آن‌ها مثبت بوده و الگوریتم دسته‌بندی نیز دسته آن‌ها را به درستی مثبت تشخیص داده است.

مهم‌ترین معیار برای تعیین کارآیی یک الگوریتم دسته بندی معیار دقت یا نرخ دسته‌بندی است. این معیار دقت کل یک الگوریتم دسته‌بندی را محاسبه می‌نماید. این معیار نشان‌دهنده این حقیقت است که دسته‌بندی طراحی شده چند درصد از کل مجموعه رکوردهای آزمایشی را به درستی دسته‌بندی کرده است. دقت دسته‌بندی با توجه به رابطه ۶ محاسبه می‌شود:

$$CA = \frac{TN + TP}{TN + FN + TP + FP} \quad (۶)$$

با توجه به رابطه بالا مقادیر TP و TN مهم‌ترین مقادیری هستند که در یک مساله دو دسته‌ای باید بیشینه شوند. در مسائل چند دسته‌ای مقادیر قرار گرفته بر روی قطر اصلی ماتریس درهم‌ریختگی باید در صورت کسر محاسبه CA قرار بگیرند

معیار Recallx: دقت دسته‌بندی دسته X را با توجه به کل رکوردهای با برچسب X نشان می‌دهد.

معیار Percisionx: دقت دسته‌بندی دسته X را با توجه به کل مواردی نشان می‌دهد که برچسب X برای رکورد موردبررسی توسط دسته‌بندی پیشنهاد شده است.

معیار Recall کارآیی دسته‌بندی را با توجه به تعداد رخداد دسته X نشان می‌دهد و معیار Precision مبتنی بر دقت پیش‌بینی دسته‌بندی می‌باشد و بیان‌گر این است که به چه میزان می‌توان به خروجی دسته‌بند اعتماد کرد.

منظور از Recall- معیار Recall برای دسته منفی است و منظور از Precision- هم معیار Precision برای دسته منفی است. این روابط قابل تعمیم برای مسائل چند دسته‌ای می‌باشند.

$$Recall^{-} = \frac{TN}{FP + TN} \quad (۷)$$

$$Percision^{-} = \frac{TN}{FN + TN} \quad (۸)$$

$$\text{Recall}^{+} = \frac{TP}{FN + TP} \quad (۹)$$

$$\text{Percision}^{+} = \frac{TP}{TP + FP} \quad (۱۰)$$

بر اساس نتایج مدلسازی، روش بیزین با دقت بیش از ۷۰٪ می‌تواند بروز تاخیر را شناسایی کند. میزان دقت در پیش بینی بازه‌های مختلف تاخیر متفاوت است. اما روش توسعه یافته، به طور متوسط در پیش بینی ۶۹،۸۱٪ میزان تاخیرها موفق بوده است. نتیجه دقت روش بیزین در پیش بینی تاخیرها در جدول ۳ آورده شده است. نتایج محاسبات مربوط به ماتریس درهم ریختگی در جدول (۴) آورده شده است. بر این اساس، دقت و قابلیت اعتماد روش پیشنهادی در تاخیرهای بیشتر از ۶۰ دقیقه بالاتر است. هر چند دقت روش در پیش بینی تاخیرهای کمتر از ۱۵ دقیقه و بین ۱۵ دقیقه تا یک ساعت هم قابل قبول است.

۶- جمع بندی و نتیجه گیری

برنامه ریزی پرواز به دلایل مختلفی از جمله نقص فنی، تاخیر در پرواز قبلی، آب و هوای نامناسب، در اختیار نبود خدمه و ... دچار اختلال می‌شود. هر نوع اختلال و تاخیر در عملیات یک شبکه پروازی به سایر پروازها انتقال و سرایت می‌یابد و مسافران و خدمه سایر پروازها را تحت تاثیر قرار می‌دهد. انتشار تاخیرها در شبکه پروازی بسیار نامطلوب است و ضررهای مالی و معنوی زیادی را برای شرکتهای هواپیمایی به همراه دارد. بنابراین محققان و پژوهشگران تلاش بسیار زیادی برای کاهش تاخیرهای تحمیلی به پروازها داشته اند. اولین گام برای کاهش این تاخیرها، پیش بینی کردن آنها در شبکه می‌باشد. در صورت پیش بینی تاخیرها می‌توان برنامه پروازی را به نحوی قابل انعطاف و مقاوم در مقابل تاخیرها بازطراحی نمود. در این مقاله ضمن بررسی پارامترهای موثر در تاخیر پروازها یک مدل داده کاوی (بیزین) برای پیش بینی تاخیر پروازها ارائه شد. مدل پیشنهادی بر روی داده‌های یکی از شرکتهای هواپیمایی بزرگ مقیاس داخلی پیاده شده است و بر اساس نتایج، مدل پیشنهادی قدرت پیش بینی بیش از ۷۰٪ وقوع تاخیرها و میزان تاخیرها را دارد.

۷- مراجع

[1] Jingqiao Zhang and Arthur C. Sanderson, “Adaptive Differential Evolution: A Robust Approach to Multimodal Problem Optimization”, Springer-Verlag Berlin Heidelberg, 2009.

[2] Guy Desaulniers, Jacques Desrosiers, Marius M. Solomon, “Column Generation”, © 2005 by Springer Science+Business Media, Inc., 2005

[3] Massoud Bazargan, “Airline Operations and Scheduling Second Edition”, Printed and bound in Great Britain by MPG book Group, UK, 2010.

[4] J. Cordeau, G. Stojkovi´c, F. Soumis, and J. Desrosiers. Benders decomposition for simultaneous aircraft routing and crew scheduling. Transportation Science, 35(4).

[5] A. M. Cohn and C. Barnhart. Improving crew scheduling by incorporating key maintenance routing decisions. Operations Research, 51:387–396, 2003.

[6] D. Klabjan, E. Johnson, G. Nemhauser, E. Gelman, and S. Ramaswamy. Airline crew scheduling with time windows and plane-count constraints. Transportation Science, 36(3):337–348, 2002.

[7] M. Lohatepanont and C. Barnhart. Airline schedule planning: Integrated models and algorithms for schedule design and fleet assignment. Transportation Science, 38(1):19–32, 2004.

[8] B. Rexing, C. Barnhart, T. Kniker, A. Jarrah, and N. Krishnamurthy. Airline fleet assignment with time windows. Transportation Science, 34(1):120, 2000.

[9] M. Stojkovi´c and F. Soumis. An optimization model for the simultaneous operational flight and pilot scheduling problem. Management Science, 47(9):1290–1305, 2001.

[10] R. Sandhu and D. Klabjan. Integrated airline fleeting and crew-pairing decisions. Operation Research, 55(3):439–456, 2007.

[11] N. Kohl, A. Larsen, J. Larsen, A. Ross, and S. Tiourine. Airline disruption management perspectives, experiences and outlook. Journal of Air Transport Management, 13:149–162, 2007.

[12] S. Bratu and C. Barnhart. Flight operations recovery - new approaches considering passenger recovery. Journal of Scheduling, 9:279–298, 2006.

[13] C. Barnhart and G. Laporte (Eds.), “Handbook in

Research 185 (2008) 825–848.

[15] Wenkai Li, Mark Wallace, “Disruption Management for Commercial Aviation”, Economics & Management Series, EMS-2012-18

Operation Research, Vol. 14”, 2007 Elsevier.

[14] Khaled F. Abdelghany, Ahmed F. Abdelghany, Goutham Ekollu, “An integrated decision support tool for airlines schedule recovery during irregular operations”, European Journal of Operational

Flight delay prediction model

Abdolreza Sheikholeslami, Hassan Khaksar, Ali Faghany

1- Associate Professor, Dept. of Civil Engineering, Iran University of Science and Technology
2- PHD Candidate, Transportation Engineering, Dept. of Civil Engineering, Iran University of Science and Technology
3- Master of Science, Transportation Engineering, Dept. of Civil Engineering, Iran University of Science and Technology

Abstract
Flights in around the world have encountered with delays usually. Such delays have many costs for airlines, passengers and society. So identifying and predicting the delays is very important problem. Purpose of this research is to predict airline delays and identify the most important factors which generate delays in airline operation. A new data mining approach has been used for delay prediction. This approach can predict occurrence a delay and amount of each delay in the flight network. Proposed methodology was implemented on a real data set of a large Iranian airline with more than 50 airplanes that fly between 52 airports. Results show that the age of fleet, aircraft type and departure time have the most effect on delays. The accuracy of approach is 70% at predicting a delay event and predicting the amount of delay.
key words : flight delay/ ostovar model / flight planing